INTERNATIONAL STANDARD

ISO 22274

First edition
2013-01-15

Systems to manage terminology, knowledge and content — Concept-related aspects for developing and internationalizing classification systems

Systèmes de gestion de la terminologie, de la connaissance et du contenu — Aspects conceptuels du développement et de la localization des systèmes des classement

Reference number
ISO 22274:2013(E)

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 22274 was prepared by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 3, *Systems to manage terminology, knowledge and content*.

Introduction

Classifying things is a common technique humans use to cope with the complexity of the world around us. The role of classification systems in our daily life can hardly be overestimated.

Classification systems organize content in a systematic way. They are highly influenced by their respective domain-specific terminologies and can, in turn, have an effect on those domain-specific terminologies. Classification systems make domain knowledge accessible to a broad audience beyond the specialists who are directly involved in that domain. Terms are established and knowledge is systematized in classification systems.

In many cases, classification systems are used to structure large collections of data supporting functions such as data mining or information retrieval. Dictionaries, libraries or catalogues, as well as web pages or retrieval systems, are examples of data collections that may benefit from being structured by classification systems.

Classification systems allow people to communicate about topics by providing sets of concepts that help to reduce the complexity of the topic to a level which is manageable for their users. These concepts allow us to direct the information flow within or between software applications, to communicate with experts from different domains or to communicate with people of different backgrounds.

If the classification system is to be used in more than one linguistic community, it needs to be localized to account for the languages, social conventions, and cultures of its users. To facilitate localization, the classification system needs to be designed so that it is clear, easy to use, and otherwise prepared to be localized.

This International Standard provides advice on how to design classification systems and how to express their content so that they are adaptable to different linguistic environments. This International Standard complements existing documents, e.g. ISO/IEC˚Guide˚77,[19] ISO 13584,[9] IEC 61360,[18] ISO 22745,[15] and ISO/IEC 11179.[7]

Systems to manage terminology, knowledge and content — Concept-related aspects for developing and internationalizing classification systems

1 Scope

This International Standard establishes basic principles and requirements for ensuring that classification systems are suitable for worldwide application, considering such aspects as cultural and linguistic diversity as well as market requirements. By applying principles relating to terminology work, this International Standard provides guidelines for creating, handling, and using classification systems for international environments.

This International Standard addresses the need in many domains for classification systems that are concept based to ensure that they are suitable for worldwide use and can be adapted to specific user communities. It provides information about the design, development, and use of classification systems that are fully enabled for diverse linguistic, cultural, and market-based environments.

This International Standard primarily specifies the factors that need to be considered when creating and populating a classification system for use in diverse linguistic environments. These factors include the specification of principles for incorporating internationalization aspects into classification systems, and maintaining and using those aspects for the structuring of activities, products, services, agents, and other entities of a company or organization.

The following are within the scope of this International Standard:

a) guidelines on information content to support internationalization of classification systems and their underlying concept systems;

b) terminological principles applicable to classification systems;

c) requirements for internationalization of classification systems;

d) considerations on workflow and administration of classification system content to support worldwide use.

The following are outside the scope of this International Standard:

— providing formal data models for representing classification systems in machine-readable form;

— prescribing classification system content for specific business domains or products;

— harmonization of classification systems.

This International Standard is intended for those who develop content for classification systems. This includes terminologists and content managers who are called upon to apply the principles of terminology work to ensure that cultural and linguistic diversity are appropriately reflected in classification systems. It is also relevant for people who design and model appropriate IT tools.

NOTE Formal data models for implementation of classification systems in information technology environments can be obtained from technical committees such as ISO/TC 184 or IEC/TC 3.

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

ISO/IEC 6523 (all parts), *Information technology — Structure for the identification of organizations and organization parts*

ISO/IEC 15418, *Information technology — Automatic identification and data capture techniques — GS1 Application Identifiers and ASC MH10 Data Identifiers and maintenance*

ISO/IEC 15459-6, *Information technology — Automatic identification and data capture techniques — Unique identification — Part 6: Groupings*

ISO/TS 29002-5, *Industrial automation systems and integration — Exchange of characteristic data — Part 5: Identification scheme*

ISO/IEC Directives, Supplement:2012, *Procedures specific to IEC*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 1087-1 and the following apply.

3.1
associative relation
relation between two **concepts** (3.7) having a non-hierarchical thematic connection by virtue of experience

EXAMPLE An associative relation exists between the concepts "education" and "teaching" or "baking" and "oven".

[SOURCE: ISO 1087-1:2000, 3.2.23, modified]

3.2
attribute
data element for the computer-sensible description of a **property** (3.25), a relation or a **class** (3.4)

[SOURCE: ISO/IEC Guide 77-2:2008, 2.2]

EXAMPLE Creation date of a class **object** (3.22) in a computer system.

3.3
characteristic
distinguishing feature

NOTE 1 A characteristic can be inherent or assigned.

NOTE 2 A characteristic can be qualitative or quantitative.

NOTE 3 There are various **classes** (3.4) of characteristic, such as the following:

— physical (e.g. mechanical, electrical, chemical or biological characteristics);

— sensory (e.g. related to smell, touch, taste, sight, hearing);

— behavioural (e.g. courtesy, honesty, veracity);

— temporal (e.g. punctuality, reliability, availability);

— ergonomic (e.g. physiological characteristic or related to human safety);

— functional (e.g. maximum speed of an aircraft).

[SOURCE: ISO 9000:2005, 3.5.1]

NOTE 4 Characteristics that apply to **concepts** (3.7) are called **features** (3.12), whereas characteristics of **classes** (3.4) are called **properties** (3.25).

EXAMPLE Figure 1 shows the interrelation of the items **concept** (3.7), **feature** (3.12), **class** (3.4), **attribute** (3.2), and **property** (3.25). The figure shows sections of a **concept system** (3.8) and a **classification system** (3.6). The class "Car" is derived from the concept "Motor vehicle" and the property "Colour" implements the feature "Pigmentation". The class "Car" has attributes such as "Preferred name" and "Definition". The concept "Automobile" is not used in the classification system.

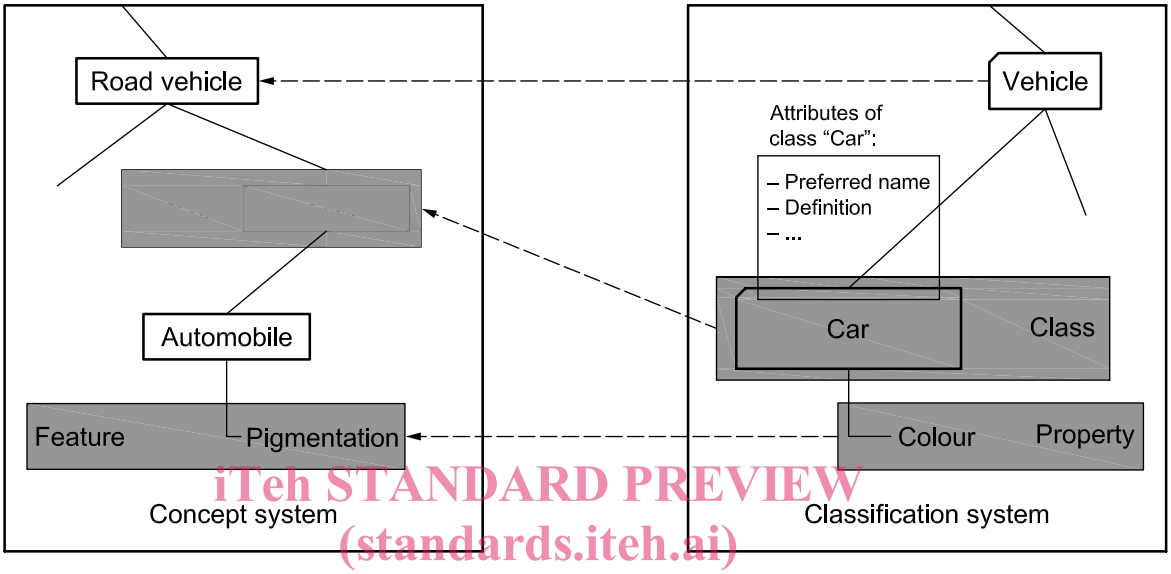


Figure 1 — Interrelation of concept (3.7), feature (3.12), class (3.4), attribute (3.2), and property (3.25)

3.4
class
description of a set of **objects** (3.22) that share the same **characteristics** (3.3)

NOTE The characteristics may be embodied by the use of properties, operations, methods, relations, semantics, etc.

3.5
classification
process of assigning **objects** (3.22) to **classes** (3.4) according to criteria

3.6
classification system
systematic collection of **classes** (3.4) organized according to a known set of rules, and into which **objects** (3.22) may be grouped

NOTE This International Standard considers both classification systems with properties and classification systems without properties.

EXAMPLE 1 The United Nations Standard Products and Services Code (UNSPSC) is an example of a classification system without properties.

EXAMPLE 2 IEC 61360-4-DB[18] is an example of a classification system with properties.

3.7
concept
unit of knowledge created by a unique combination of **characteristics** (3.3)

[SOURCE: ISO 1087-1:2000, 3.2.1]

NOTE Concepts are not necessarily bound to particular languages. They are, however, influenced by the social or cultural background which often leads to different **classifications** (3.5).

3.8
concept system
set of **concepts** (3.7) structured according to the relations among them

[SOURCE: ISO 1087-1:2000, 3.2.11]

3.9
extension
totality of **objects** (3.22) to which a **concept** (3.7) corresponds

[SOURCE: ISO 1087-1:2000, 3.2.8]

3.10
facet
group of **classes** (3.4) or **concepts** (3.7) of the same inherent category

[SOURCE: ISO 25964-2:—,[16] 3.32, modified]

EXAMPLE 1 High-level categories that can be used for grouping concepts into facets are: **objects** (3.22), materials, agents, actions, places and items.

NOTE Facets used in **classification systems** (3.6) should follow the rules given in 5.4, whereas facets used in **concept systems** (3.8) are free from such restrictions. In either case, the recommendations given in 5.3 should apply.

EXAMPLE 2 Facets of a **classification system** (3.6) for commodities may be functional view, product-oriented view, material, maintenance considerations or logistics.

3.11
faceted classification system
classification system (3.6) where **classes** (3.4) are grouped in mutually exclusive and collectively exhaustive aspects that can be combined to specify complex subjects

EXAMPLE Classes to specify programmable logic controllers may be grouped in **facets** (3.10) such as “technology”, “programming”, “packaging”, and “accounting”.

3.12
feature
defined **characteristic** (3.3) suitable for the description and differentiation of **concepts** (3.7) in a **concept system** (3.8)

3.13
general concept
concept (3.7) which corresponds to two or more **objects** (3.22) which form a group by reason of common **characteristics** (3.3)

EXAMPLE “Planet” or “tower”.

[SOURCE: ISO 1087-1:2000, 3.2.3, modified]

3.14
generic relation
relation between two **concepts** (3.7) where the **intension** (3.15) of one of the **concepts** (3.7) includes that of the other **concept** (3.7) and at least one additional delimiting **characteristic** (3.3)

[SOURCE: ISO 1087-1:2000, 3.2.21, modified]

3.15
intension
set of **characteristics** (3.3) of a **concept** (3.7)

[SOURCE: ISO 1087-1:2000, 3.2.9, modified]

3.16
internationalization
process whereby products and services are implemented in a way that allows for and facilitates the adaptation to local languages and cultural conventions

NOTE Internationalization is a prerequisite for a systematic approach to **localization** (3.21).

3.17
leading locale
locale (3.20) in or for which a product or service is developed and which serves as a reference point for further **localization** (3.21)

3.18
leaf class
class (3.4) in a hierarchical **classification system** (3.6) which has one or more superordinate classes and no subordinates

3.19
level
magnitude of a quantity considered in relation to a reference value

3.20
locale
unique combination of parameters specifying the language, geographic area, and other cultural, administrative or technical preferences of a given community

3.21
localization
adaptation of a product or communication to a community of speakers with respect to cultural, linguistic, legal, political and technological factors

[SOURCE: ISO/TR 22134:2007,[14] 3.7]

3.22
object
anything perceivable or conceivable

NOTE Objects may be material (e.g. an engine, a sheet of paper, a diamond), immaterial (e.g. conversion ratio, a project plan) or imagined (e.g. a unicorn).

[SOURCE: ISO 1087-1:2000, 3.1.1]

3.23
partitive relation
relation between two **concepts** (3.7) where one of the concepts constitutes the whole and the other concept a part of that whole

NOTE A partitive relation exists between the concepts “week” and “day” or “molecule” and “atom”.

[SOURCE: ISO 1087-1:2000, 3.2.22, modified]

3.24
polyhierarchy
hierarchy including elements that have links to more than one parent element

3.25
property
defined **characteristic** (3.3) suitable for the description and differentiation of the **objects** (3.22) in a **class** (3.4)

EXAMPLE Ambient temperature may be a property of a class comprising geographical locations.

3.26
terminology
set of designations belonging to one special language

[SOURCE: ISO 1087-1:2000, 3.5.1]

3.27
value domain
set of permissible values

[SOURCE: ISO/IEC 11179-1:2004,[7] 3.3.38]

3.28
vocabulary
terminological dictionary which contains designations and definitions from one or more specific subject fields

[SOURCE: ISO 1087-1:2000, 3.7.2]

4 Relations to other documents

The following documents provide guidelines about fundamentals applicable to the development of classification systems and to other related tasks.

— ISO 704[1] defines the essential elements for quality in terminology work.

— ISO 1087-1 defines fundamental terminological concepts.

— ISO/TS 29002-5 specifies elements and syntax of identifiers of elements of a concept dictionary.

— IEC 61360,[18] ISO 13584[9] or ISO 22745[15] specify data models that may be used to store, retrieve or maintain classification systems in data-processing environments.

— Annex SL, *Procedures for the maintenance of the IEC standards in database format*, in: ISO/IEC Directives, Supplement:2012, describes procedures applicable for the maintenance of International Standards comprising classification systems managed in a data-processing environment.

5 Fundamental development considerations for classification systems

5.1 General

Classification systems are widely used to facilitate handling and interpretation of objects by organizing the knowledge of an area of discourse and thus provide the information needed to unambiguously characterize those objects. To serve this purpose, classification systems should be carefully designed to avoid structures that do not provide the requested information or that are overly complicated, which confuses users. The development of a sound classification hierarchy is facilitated if it reflects an underlying concept system that shall be elaborated based on recognized principles for managing terminology. Through a mapping process, the concepts in the concept system become classes in the classification system.

Often the criterion of classification is likeness. A classification unites like things and it separates unlike things. Things may, however, be alike in many different ways. A classification should unite things from

a functional or a pragmatic point of view based on the purpose of the classification. The objects to be classified may be things, persons, processes, ideas, services, and so forth.

5.2 Application domains

Classification systems are developed for many domains. They are used in domains such as:

— healthcare;

— manufacturing;

— service delivery;

— documents and libraries;

— science;

— retail.

Examples 1 to 8 describe sample classification systems from each of these domains.[1)]

EXAMPLE 1 The Global Medical Device Nomenclature (GMDN) is an important classification system for medical devices. All parties involved with medical devices, such as manufacturers, regulators, conformity assessment bodies, traders, owners, and users, have a common interest in having access to an unambiguous classification of those devices, including clear definitions and terms. Processes addressed by GMDN include:

— manufacturing;

— registration;

— incident reporting;

— trading;

— inventory, stock-keeping, and life-cycle information.

EXAMPLE 2 eCl@ss® (Reference [32]) is an international industry standard for classification of products, materials, and services. In addition to providing the option of describing each individual product in a unique way, eCl@ss® provides codes to specify product groups. Purchasers and consumers can identify products and services with the eCl@ss® codes when they are using product databases, merchandise information systems or electronic catalogues.

EXAMPLE 3 The International Standard Industrial Classification (ISIC) is an international reference classification of activities within production processes. Its main purpose is to provide a set of activity categories that can be utilized for the collection and reporting of statistics according to such activities. Since the original version in 1948, ISIC has provided guidance to countries for developing national activity classifications and has become an important tool for comparing statistical data on economic activities at the international level.

EXAMPLE 4 The Universal Decimal Classification (UDC) is a multilingual classification scheme for all fields of knowledge, a sophisticated indexing and retrieval tool for organizing bibliographic records for all kinds of information in any medium. It was adapted from the Dewey Decimal Classification (DDC) and first published (in French) between 1904 and 1907.

EXAMPLE 5 The London Classification of Business Studies (LCBS) was compiled as a result of rapid expansion in the field of management education following the establishment of two graduate business schools in the UK (London and Manchester) in 1965. It has an international reputation and is used in many business libraries and information services throughout the world.

EXAMPLE 6 The Library of Congress Classification (LCC) is a classification system that was first developed in the late 19th and early 20th century to organize the book collections of the Library of Congress. Over the course of the 20th century, the system was adopted for use by other libraries as well, especially large academic libraries in the USA and in research and academic libraries in other countries.

1) The Global Medical Device Nomenclature (GMDN), eCl@ss®, The Physics and Astronomy Classification Scheme® (PACS®), GS1®, and GSDN® are examples of suitable services available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of these services.

EXAMPLE 7 The Physics and Astronomy Classification Scheme® (PACS®) is designed to categorize the literature of physics and astronomy. PACS® is used by international publishers of journals in physics, astronomy, and related fields.

EXAMPLE 8 The Global Product Classification (GPC) is the classification system of GS1® (Reference [22]) representing users in more than 20 activity sectors. GPC currently classifies 36 product category segments. GPC is mandatory in the Global Data Synchronisation Network (GDSN®) and fully in line with the GS1® system.

5.3 Generic requirements

5.3.1 Consistency

A key requirement on classification systems is consistency. Classes shall be clearly separated from each other and their individual areas of applicability shall not overlap. Especially in cases when processing by information technology is envisaged, a clear structure and the absence of ambiguities is a requirement. Definitions shall clearly identify the specific concepts that make up the various classes.

The terms used to name the classes as well as those used in the definitions of classes and any other information in the classification system shall be consistent. A consistent terminology provides the basis for an unambiguous communication between users as well as between software applications.

5.3.2 Comprehensibility

The rules for creating classes and for writing definitions shall be documented and made available for users and providers of content. Clear definitions of the concepts employed in the classification system make it more comprehensible and reduce ambiguities. The levels of the classification system and their divisions into individual classes shall be explicitly defined. The criteria or rules for creating new hierarchy levels or new classes shall be documented.

5.3.3 Extensibility

Extensibility is the ability to accommodate new classes at their correct place in the classification system. Classes shall be created in a level or position in the hierarchy that reflects their relations to existing classes. Classification systems shall be extensible so that they can accommodate new requirements or perspectives within their area of applicability.

5.4 Structuring principles

5.4.1 General

Even though most of the existing classification systems do not exclusively apply one of the approaches described below, but are a blend of them, the following main principles can be identified:

— enumerative;

— faceted;

— enumerative and faceted (with entry class).

There is no generic rule on the structure of a classification system. The decision has to be made on a case-by-case basis in accordance with the requirements that result from the intended area of application. This includes subjects such as:

— structuring principle;

— use of properties (mandatory or optional);

— properties at any level or properties at leaf level only.

Additionally, changes of already existing structures may become necessary during the lifetime of a classification system.

5.4.2 Enumerative classification systems

Enumerative classification systems attempt to list all possible subjects within their defined area of applicability. They are in many cases represented by using hierarchies. Nevertheless, in some cases enumerative schemes may be represented by simple unstructured sets of objects.

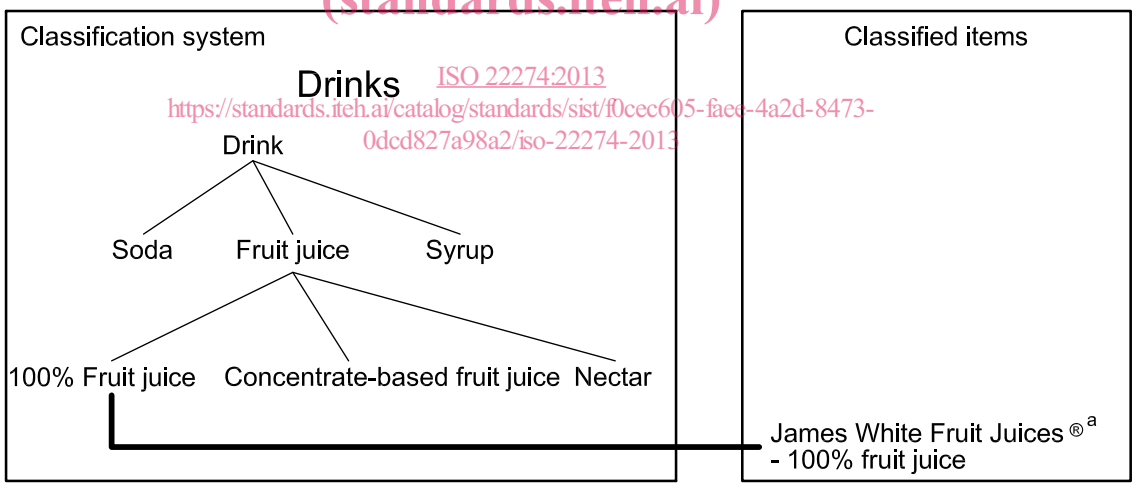
NOTE 1 See Reference [26] for an in-depth treatment of enumerative classification systems.

The hierarchy utilizes a "top down" approach, a process of division producing a series of classes in successive subordination. Thus, the number of subclasses of any class may require a limitation for ease of handling by the intended users of the classification system.

NOTE 2 Successive subordination of classes creates a hierarchical structure of the classification system. Thus, enumerative classification systems are also referred to as hierarchical classification systems.

NOTE 3 In enumerative classification systems, a subject can only be classified if it is explicitly covered by the area of applicability of one of the classes in the classification system. Enumerative classification systems can, therefore, require a higher number of classes than faceted classification systems, because for each class all possible combinations of constituent characteristics require representation. Additionally, it is possible for constituent characteristics such as "unfinished" to require frequently repetition. Thus, an enumerative classification system tends to become a compromise between the number of its classes and its completeness. A complete enumerative classification system is often very complex in nature and its basic principles of construction are difficult to identify.

NOTE 4 Such classification systems often include general classes such as "Miscellaneous" for classes that do not precisely fit into a more specifically named class.



Key

——— Generic relation

NOTE This figure is for illustrative purposes only. It is not intended to be a valid classification system for the domain of drinks.

[a] Product available commercially. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of this product.

Figure 2 — Example of an enumerative classification system

5.4.3 Faceted classification systems

Faceted classification systems allow the assignment of multiple classifications to an object. An object may be characterized by any combination of the classes from the facets.