



DPC: 25/30472169 DC

BSI Group Headquarters

389 Chiswick High Road London W4 4AL

Tel: +44 (0)20 8996 9000

Fax: +44 (0)20 8996 7400

www.bsigroup.com

Date: 19 May 2025

Origin: International

Latest date for receipt of comments: 12 July 2025

Project No. 2023/00748

Responsible committee: IDT/2/18 Identifiers and metadata

Interested committees:

Title: Draft BS ISO 25964-1 Information and documentation - Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval; management and use.

Please notify admin.start@bsigroup.com if you are aware of any keywords that might assist in clarifying or identifying the standard or of the content of this standard:

- i) has any issues related to 3rd party IPR, patent or copyright
- ii) affects other national standard(s)
- iii) requires additional national guidance or information

WARNING: THIS IS A DRAFT AND MUST NOT BE REGARDED OR USED AS A BRITISH STANDARD. THIS DRAFT IS NOT CURRENT BEYOND 12 July 2025

This draft is issued to allow comments from interested parties; all comments will be given consideration prior to publication. No acknowledgement will normally be sent. **See overleaf for information on the submission of comments.**

No copying is allowed, in any form, without prior written permission from BSI except as permitted under the Copyright, Designs and Patent Act 1988 or for circulation within a nominating organization for briefing purposes. Electronic circulation is limited to dissemination by e-mail within such an organization by committee members.

Further copies of this draft may be purchased from BSI Shop <http://shop.bsigroup.com> or from BSI Customer Services, Tel: +44(0) 20 8996 9001 or email cservices@bsigroup.com. British, International and foreign standards are also available from BSI Customer Services.

Information on the co-operating organizations represented on the committees referenced above may be obtained from <https://standardsdevelopment.bsigroup.com/>.

E-mail: admin.start@bsigroup.com

Introduction

This draft standard is based on national and international discussions. Your comments on this draft are invited and will assist in the preparation of the consequent standard.

For international standards, comments will be reviewed by the relevant UK national committee before sending the consensus UK vote and comments to the international committee, which will then decide appropriate action. If the international standard is approved, it is usual for the text to be published as a British Standard.

For national standards, comments will be reviewed by the relevant UK national committee and the resulting standards published as a British Standard.

UK Vote

Please indicate whether you consider the UK should submit a negative (with supporting technical reasons) or positive vote on this draft. Please indicate if you are aware of any reason why this draft standard should not be published as a British Standard.

Submission of Comments

- **Annotated drafts are not acceptable and will be rejected.**
- All comments should be submitted online at <https://standardsdevelopment.bsigroup.com/>. You will need to register in order to comment.

Template for comments and secretariat observations

Date: xx/xx/20xx	Document: ISO/DIS xxxx
------------------	------------------------

1	2	(3)	4	5	(6)	7
MB	Clause No./ Subclause No./Annex (e.g. 3.1)	Paragraph/ Figure/Table/Note	Type of comment	Comment (justification for change) by the MB	Proposed change by the MB	Secretariat observations on each comment submitted
	3.1	Definition 1	ed	Definition is ambiguous and needs clarifying.	Amend to read '...so that the mains connector to which no connection...'	
	6.4	Paragraph 2	te	The use of the UV photometer as an alternative cannot be supported as serious problems have been encountered in its use in the UK.	Delete reference to UV photometer.	



DRAFT International Standard

ISO/DIS 25964-1

Information and documentation — Thesauri and interoperability with other vocabularies —

Part 1: Thesauri for information retrieval; management and use

ICS: 01.140.20

ISO/TC 46/SC 9

Secretariat: **ANSI**

Voting begins on:
2025-05-19

Voting terminates on:
2025-08-11

This document is circulated as received from the committee secretariat.

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENTS AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

Reference number
ISO/DIS 25964-1:2025(en)

Full document details: 25/30472169 DC

© ISO 2025  **Documents**
Document Sharing Hub



COPYRIGHT PROTECTED DOCUMENT

© ISO 2025

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
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	vii
Introduction	viii
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Symbols, abbreviated terms and other conventions	21
5 Thesaurus overview and objectives	25
5.1 Overall objective	25
5.2 Vocabulary control and its purpose	26
5.3 Paradigmatic versus syntagmatic relationships	26
5.4 Types of paradigmatic relationship	27
6 Concepts and their scope in a thesaurus	28
6.1 Conceptual basis	28
6.2 Scope notes	30
6.3 Reciprocal scope notes	31
7 Thesaurus terms	32
7.1 Form of terms	32
7.2 Term attributes	32
7.3 Selection of the preferred term for display as concept label	32
7.3.1 General	32
7.3.2 Spellings	32
7.3.3 Loan terms and translations of loan terms	34
7.3.4 Transliteration	34
7.3.5 Neologisms, slang terms and jargon	34
7.3.6 Common names and trade names	35
7.3.7 Popular names and scientific names	35
7.3.8 Abbreviations and acronyms	35
7.3.9 Proper names	36
7.4 Clarification and disambiguation of thesaurus terms	38
7.4.1 General	38
7.4.2 Homographs and qualifiers	38
7.4.3 Definitions	39
7.4.4 History notes	40
7.5 Grammatical form of terms	40
7.5.1 Nouns and noun phrases	40
7.5.2 Adjectives	41
7.5.3 Adverbs	41
7.5.4 Verbs	41
7.5.5 Initial articles	42
7.6 Capitalization, punctuation and special characters	42
7.6.1 Capitalization	42
7.6.2 Non-alphabetic characters	43
7.7 Singular or plural forms	44
7.7.1 Cultural and linguistic factors	44
7.7.2 Treatment of count nouns in English and Spanish	44
7.7.3 Treatment of non-count nouns in English and Spanish	45
7.7.4 Coexistence of singular and plural	45
8 Complex concepts	46
8.1 General	46
8.2 The nature of compound terms	47
8.3 Deciding whether or not to admit a complex concept	48
8.3.1 The options in outline	48

8.3.2	Factors to consider	48
8.3.3	Circumstances that favour splitting a complex concept.....	49
8.3.4	Circumstances that rule against splitting a complex concept.....	50
8.4	How to split a complex concept.....	52
8.5	Retention of constituent concepts.....	53
8.5.1	General	53
8.5.2	Parts and components.....	53
8.6	Consistency in the treatment of complex concepts	53
8.7	Order of words in multi-word terms.....	54
9	The equivalence relationship, in a monolingual context.....	54
9.1	General.....	54
9.2	Synonyms.....	55
9.3	Quasi-synonyms.....	58
9.4	Specific terms subsumed in a broader concept.....	59
10	Equivalence across languages	59
10.1	General.....	59
10.2	Degrees of equivalence	60
10.3	Typical problems and solutions in multilingual thesauri.....	62
10.3.1	General	62
10.3.2	Problems caused by quasi-synonyms or homographs	62
10.3.3	Absence of an equivalent term in one or more languages.....	64
10.3.4	Combined problems.....	66
10.4	Representation of cross-language equivalence between preferred terms.....	67
10.5	Cross-language equivalence between non-preferred terms	67
11	Relationships between concepts	68
11.1	Introduction.....	68
11.2	The hierarchical relationship.....	68
11.2.1	General	68
11.2.2	The generic relationship.....	69
11.2.3	The hierarchical whole-part relationship.....	71
11.2.4	The instance relationship.....	72
11.2.5	Polyhierarchical relationships.....	73
11.2.6	Non-symmetrical structures in multilingual and multicultural thesauri.....	73
11.3	The associative relationship.....	74
11.3.1	General	74
11.3.2	Terms and concepts with overlapping meanings.....	74
11.3.3	Other cases of associative linkage.....	75
11.3.4	Multilingual and multicultural thesauri.....	78
11.4	Customized relationships	78
12	Facet analysis.....	79
13	Presentation and layout.....	82
13.1	General.....	82
13.1.1	Use of a database for maintaining presentation information.....	82
13.1.2	Summary of display styles.....	82
13.1.3	Notation and links between sequences	82
13.2	Alternative display styles	83
13.2.1	General	83
13.2.2	Single record display	83
13.2.3	Alphabetical display.....	84
13.2.4	Hierarchical display.....	86
13.2.5	Classified display	88
13.2.6	Displaying polyhierarchical relationships	92
13.3	Presentation and layout of multilingual thesauri	92
13.3.1	General.....	92
13.3.2	Single-record display.....	92
13.3.3	Alphabetical display.....	93

13.3.4	Systematic displays.....	94
13.3.5	Correspondence tables for multilingual thesauri.....	96
13.4	Language and character encoding issues.....	98
13.4.1	General.....	98
13.4.2	Display issues.....	99
13.4.3	Filing orders.....	99
13.4.4	Normalization for information retrieval.....	100
13.4.5	Identification of dialects and/or alternative scripts.....	100
13.5	Concept identifiers and resource locators for use in controlled vocabularies.....	101
14	Managing thesaurus construction and maintenance.....	102
14.1	Planning a thesaurus.....	102
14.1.1	Determining objectives.....	102
14.1.2	Determining thesaurus features.....	102
14.1.3	Taking stock of resources.....	103
14.1.4	Establishing responsibilities.....	104
14.1.5	Choosing thesaurus management software.....	104
14.2	Early stages of compilation.....	104
14.2.1	General: how and when to start.....	104
14.2.2	Collecting terms and concepts.....	104
14.2.3	Analysing terms.....	105
14.3	Construction.....	105
14.3.1	Input.....	105
14.3.2	Construction sequence.....	105
14.3.3	Construction of a multilingual thesaurus.....	106
14.3.4	Involving experts and stakeholders.....	107
14.3.5	Appropriate choice of terminology.....	107
14.4	Introduction to the thesaurus.....	108
14.5	Dissemination.....	108
14.5.1	Integration with an electronic system.....	108
14.5.2	Electronic publication.....	109
14.5.3	Limited distribution of printed copies.....	109
14.5.4	Conventional publication.....	109
14.5.5	Deposit with a clearinghouse.....	109
14.5.6	Website directories.....	110
14.5.7	Updating.....	110
14.5.8	Suggestion procedure.....	110
14.5.9	Review procedure.....	110
14.5.10	The nature of changes.....	111
14.5.11	Dissemination of updates.....	112
15	Data model.....	113
15.1	General.....	113
15.2	The Visual Model.....	114
15.3	Tabular presentation.....	119
15.4	Integration of thesauri with applications.....	124
15.5	Exchange formats.....	127
15.5.1	MARC (MACHine-Readable Cataloguing).....	127
15.5.2	SKOS.....	128
15.5.3	Zthes.....	128
15.5.4	Additional markup domains.....	128
15.6	Protocols.....	128
15.6.1	General.....	128
15.6.2	Purposes and use cases.....	128
15.6.3	Application environment and architecture.....	128
15.6.4	General-purpose web database protocols used with thesauri.....	129
Annex A (informative) Examples of displays found in published thesauri.....		130
Annex B (informative) XML Schema for data exchange.....		144

Annex C (informative) Guidelines for thesaurus management software	145
Bibliography	150
Index	158

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.

This document was prepared by Technical Committee ISO/TC 46, Information and documentation, Subcommittee SC 9, Identification and description.

This second edition cancels and replaces the first edition (ISO 25964-1:2011), which has been technically revised.

The main changes are as follows:

- Updated examples and references
- Added concept and concept labels to differentiate meaning (concepts) and the representations (terms)
- Added and updated definitions and terms, for example for the following: concepts candidate concept, Lexical labels, sequence numbers, classification codes, GUIDs and UUIDs, information retrieval, information use, information management and mapping, etc. 17 additional definitions are now included
- Updated Equivalence relationships
- Updated Relationships between concepts
- Added guidance on complex concepts
- Added a section on PID
- Revised the Data Model and identifiers guidance
- Added significant number of multi-language examples
- Added more guidance on automation and artificial intelligence
- Updated the references both for currency and accuracy

A list of all parts in the ISO 25964 series can be found on the ISO website.

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

Today's thesauri are mostly electronic tools, having moved on from the paper-based era when thesaurus standards were first developed. They are built and maintained with the support of software and need to integrate with other software, such as search engines and content management systems. (For example, data from the thesaurus database might need to be presented in combination with the number of postings found by a search application.) Whereas in the past thesauri were designed for information professionals trained in indexing and searching, today there is a demand for vocabularies that untrained users will find to be intuitive, and for vocabularies that enable inferencing by machines such as semantic web, linked data, and artificial intelligence (AI) (meaning NLP, machine learning, LLM, etc.).

ISO 25964 makes the transition that is needed in order to be compatible with the world of deeply inter-linked electronic information management. However, this part of ISO 25964 retains the assumption that human intellect is usually involved in the selection of indexing terms and in the selection of search terms even while suggestions may be made using automated methods. If both the indexer and the searcher are guided towards the same concept, then relevant documents will be retrieved. This was the main principle underlying thesaurus design and still applies. However, we are increasingly in a world where a thesaurus is used for the annotation of machine generated responses, or guidance of AI by thesauri and other structured vocabularies. In search and navigation, computers increasingly make the choices of what information is presented to the searcher.

Efficient exchange of data is a vital component of thesaurus management and exploitation. This part of ISO 25964 therefore includes recommendations for exchange formats and protocols. Adoption of these will facilitate interoperability between thesaurus management systems and other computer applications, such as indexing and retrieval systems and artificial intelligence (AI) including generative language models, including generative language models that will utilize the data.

This part of ISO 25964 covers development and maintenance of thesauri rather than how to use them in indexing. Where multilingual issues and examples are addressed, efforts have been made to cover as wide a selection of languages as possible, consistent with clarity and comprehensibility.

Thesauri are typically used in post-coordinate retrieval systems, but may also be applied to hierarchical directories, pre-coordinate indexes and classification systems. Increasingly, thesaurus applications need to mesh with others, such as automatic categorization schemes, free-text search systems, etc. ISO 25964-2 addresses additional types of structured vocabulary (such as classification schemes, name authority lists, ontologies, etc.) and give recommendations to enable interoperability of the vocabularies at all stages of the information storage, retrieval, creation and management process.

In revision of the ISO 25964-1 the working group struggled with the use of Preferred ([clause 3.59](#)) and Non-preferred terms ([Clause 3.54](#)) as the term of art for the main terms applied and their equivalence terms. "Concept label" was added ([3.20](#)) to insure a clear cross walk with the W3C standards definitions and vocabularies. We have added the definitions for each area within the definition's sections providing a cross walk to other major terminology standards including the ISO 704, ISO 30042 -TBX, and ISO 1087 terminology standards as well as the definitions from the W3C SKOS and OWL communities. Many other appellations are used for both labels. Preferred term is often called default term or default concept as well as concept label, and descriptor. Non-preferred term may be referred to as Alternative term, Alternative label, AltLabel, Synonym, and others. This is not an exhaustive list for either but indicative. We encourage the community to use the term label best suited to their own community.

Information and documentation — Thesauri and interoperability with other vocabularies —

Part 1: Thesauri for information retrieval; management and use

1 Scope

Part 1 of ISO 25964 gives recommendations for the development and maintenance of thesauri intended for information retrieval, applications as well as broader use cases including web navigation, artificial intelligence, author tagging, management, and use applications. It is applicable to vocabularies used for retrieving information from all types of information resources, irrespective of the media used (text, sound, still or moving image, physical object or multimedia) including knowledge bases web sites, bibliographic databases, text, museum or multimedia collections, and the items within them, research and operational datasets, medical records, product catalogs, (see [section 3.79](#) for a definition of thesauri).

This part of ISO 25964:

- also provides a data model and recommended format for the import and export of thesaurus data.
- is applicable to monolingual and multilingual thesauri.
- is not applicable to the preparation of back-of-the-book indexes, although many of its recommendations could be useful for that purpose.
- is not applicable to the databases or software used directly in search, indexing applications as well as broader use cases including web navigation, artificial intelligence, author tagging, AI applications, but does anticipate the needs of such applications among its recommendations for thesaurus management (see also ISO 25964-2 on interoperability with other vocabularies).

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. Cross references to other ISO authorities are added into the appropriate definition.

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

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <https://www.electropedia.org/>

3.1

array

group of sibling concepts especially in a display under their shared *broader concept* ([3.3](#))

Note 1 to entry: The sibling concepts ([3.52](#)) form an array within the same immediate *broader concept* ([3.3](#)).

EXAMPLE

Example of siblings at several different levels in a hierarchy

marble by color or pattern
 ...black marble
 ...gray marble
 ...green marble
 ...variegated marble
brocatelle marble
fiorto marble
forest marble
pagoda stone
Proconnesian marble
bianco e nero antico
bigio e nero antico

3.2

associative relationship

relationship between a pair of designated *concepts* ([3.14](#)) that are not related hierarchically but share a strong semantic connection

EXAMPLE

two materials are related to each other: the stone "travertine" is distinguished from "tufa"

an implement is related to a technique: "watercolor brushes" are required for the "watercolor painting" technique

a feature is related to works: "caves" are the locus/setting for "cave paintings"

Note 1 to entry: These are also used in Authority files as well as thesauri.

3.3

broader concept

concept ([3.14](#)) that is broader than the one in question

Note 1 to entry: The scope of the narrower concept falls completely within the scope of the broader.

Note 2 to entry: The hierarchical relationship is further defined in [3.29](#).

Note 3 to entry: This is also the name of a relationships type in SKOS.

Note 4 to entry: ISO 1087 refers to this as a superordinate/broader concept.

EXAMPLE

the nation of "Kenya" is the broader context for "Lamu County," which is the broader context for "Lamu Old Town"

"dental chisels" and "dental clamps" fall completely within the scope of "dental instruments"

3.4**broader term**

preferred term ([3.56](#)) representing a *concept* ([3.14](#)) that is broader than the one in question

Note 1 to entry: The scope of the narrower concept falls completely within the scope of the broader. The relationship between the two is commonly indicated with the tag BT. For more explanation see 0. Described in hierarchical relationship [3.32](#) and that it has a reciprocal of narrower term.

3.5**candidate concept**

concept ([3.14](#)) and its terms ([3.78](#)) under consideration for inclusion in the thesaurus, but not yet approved for use.

Note 1 to entry: Also linguistic borrowing from one language to another.

See also *loan term* [3.39](#)

3.6**characteristic of division**

attribute by which a *concept* ([3.14](#)) can be subdivided into an *array* ([3.1](#)) of *narrower concepts* ([3.47](#)), each having a distinct value of that attribute

Note 1 to entry: Type of characteristic (by age group, by colour, etc) and the same concept is designated as a “criterion of subdivision” in ISO 1087, we could examine if these terms designate the same concept, i.e. if they are synonym. Note that “criterion of subdivision” is already contained in ISO 5127 (not defined) in the definitions of the concepts “extensional definition” and “coordinate concept”.

cf. facet analysis ([3.28](#)), node label ([3.50](#))

EXAMPLE

bridges may be grouped by form and by function:

bridges

...<bridges by form>

.....arch bridges

.....covered bridges

.....multispan bridges

.....truss bridges

...<bridges by function>

.....footbridges

.....military bridges

3.7**classifying**

activity involving the components of grouping similar or related things together, separating dissimilar or unrelated things and arranging the resulting groups in a logical and helpful sequence

3.8**classification**

assignment of items (e.g. documents), artifacts, etc.) into the classes of a classification scheme

3.9**classification code**

alpha-numerical identifiers that express a classification structure in notation style (see [13.5.3](#))

3.10

classification scheme

schedule (3.62) of *concepts* (3.14) and pre-coordinated combinations of concepts (3.11), arranged in a hierarchy classes and subclasses for characteristics of division (grouping things) (3.6)

Note 1 to entry: Classification schemes often include an alpha numeric notation.

Note 2 to entry: A classification scheme often also includes an index.

Note 3 to entry: Compares with ISO 1087:2019 classification scheme definition.

3.11

coined term

new term (3.78) created to express a concept (3.16) for which no suitable term exists in the required language

Note 1 to entry: Also linguistic borrowing from one language to another.

Note 2 to entry: See also loan term 7.33 .

EXAMPLE

The term for a performance art form, "happenings." was coined by the artist Allan Kaprow in the 1950.

The term for the Chinese art style. "Cynical Realism." was coined by art critic Li Xianting in the 1980s.

Note 3 to entry: For a further explanation and examples, see 10.3.3.3.

3.12

compound equivalence

relationship or mapping in which one term (3.40) or *concept* (3.14) in one context is represented by two or more terms or *concepts* (3.14) in another context or controlled vocabulary

Note 1 to entry: See also *equivalence relationship* (3.25)

3.13

compound term

term (3.77) that can be split morphologically into separate components. a compound term can be a complex term (steamships, copper mines) or a blend (e.g. biodiversity)

Note 1 to entry: Compound terms may be terms which have a one meaning when compounded but a different meaning for the terms presented individually. Also known as a bound term. That is, singularly the terms would mean something else entirely.

EXAMPLE

In English: “copper mines” can be split into “copper” and “mines”; “lawnmowers” can be split into “lawns” and “mowers”
In French: “mine de cuivre” can be split into “mine” and “cuivre”; the French plural “mines” + “cuivre” “biodiversité” can be split into “biologie” and “diversité”
In German “copper mines” in German Kupferbergwerke , Kupfer and Bergwerke. compounds in German are expressed as a single word. _
In Spanish “ minas de cobre ”
urban landscape (English), stedelijk landschap (Dutch), paisagem urbana (Portuguese), 都市景觀 (Chinese traditional)

Note 2 to entry: Compound terms can be multi-word terms or can consist of only one word.

Note 3 to entry: Compare with ISO 1087:2019 - compound term *single word term* that can be split morphologically into separate components and complex term, *term* that consists of more than one word or lexical unit.

3.14

concept

unit of thought

unit of knowledge created by a unique combination of characteristics

Note 1 to entry: Concepts can often be expressed in a variety of different ways. They exist in the mind as abstract entities independent of terms used to express them. They range from the very simple, e.g. “child”, to the very complex, e.g. “child protection legislation”. Concepts can be represented by one or more terms or labels

Note 2 to entry: Compare with ISO 1087:2019 – concept unit of knowledge created by a unique combination of characteristics, Concepts are not necessarily bound to particular natural languages. They are, however, influenced by the social or cultural background which often leads to different categorizations.

Note 3 to entry: Compare with NISO Z39.19 -2005 - A unit of thought, formed by mentally combining some or all of the characteristics of a concrete or abstract, real or imaginary object. Concepts exist in the mind as abstract entities independent of terms used to express them.

3.15

concept label

default representation of a *concept* ([3.14](#)) in a display or user interface, by a sign which denotes it in a particular domain such as a display or user interface

default designation

default designator

synonym for preferred term

c.f. *preferred term* ([3.56](#))

c.f. *identifier* ([3.31](#)) and *index term* ([3.32](#))

Also known as default designation, default designator, default representation of a *concept* (3.14) in a display or user interface, by a sign which denotes it in a particular domain.

Note 1 to entry: The preferred term is chosen as the default label for each concept in a monolingual thesaurus that uses the terminology of preferred term/nonpreferred term.

Note 2 to entry: In the case of multilingual thesauri concept label maybe used so preference for a single language is not shown.

Note 3 to entry: Described in ISO 1087:2019 as “designation, representation of a concept by a sign which denotes it in a domain or subject and term, designation that represents a general concept by linguistic means”

3.16

controlled vocabulary

prescribed list of *terms* (3.77), headings or codes, each representing a *concept* (3.14)

Note 1 to entry: Controlled vocabularies are designed for applications in which it is useful to identify each concept with one consistent label, for example when classifying documents, indexing them and/or searching them. Thesauri, subject heading schemes and name authority lists are examples of controlled vocabularies.

Note 2 to entry: Symbol languages and codes may also form a controlled vocabulary with definitions of the symbols which would form a controlled vocabulary.

3.17

cross-language equivalence

equivalence relationship (3.23) between *terms* (3.77) representing the same *concept* (3.14) in different languages

Note 1 to entry: Described in ISO 1087:2019 as “equivalence, relation between designations in different natural languages representing the same concept”.

EXAMPLE

The French term “ordinateur” is an equivalent of the English term “computer”.

English: gold; Dutch: goud; Swedish: guld; German: Gold; French: or; Italian and Spanish: oro; Portuguese: ouro; Arabic: بَهْد; Latin: aurum; symbol: Au Greek: χρυσός;

3.18

data model

abstract model that provides the representation of data, specifying their properties, structures and interrelationships

Note 1 to entry: The data model in this part of ISO 25964 provides a generic definition of thesaurus structure and semantics. It can be used as the basis for defining a database model or an exchange format for thesauri.

[SOURCE: ISO/IEC 2382-36:2019(en), 3.10.1]

3.19

deprecated term

term no longer in active use but which has been used in the past in the literature and is noted as such

Note 1 to entry: To completely remove it would not allow access to historical material.

EXAMPLE

in reference to the Pre-Columbian Southwestern North American culture concept, **Ancestral Puebloan**, a term used in past decades, “Anasazi,” is pejorative and has been deprecated, thus not recommended for use in new indexing, although it is used for retrieval of older existing materials.

Note 2 to entry: A deprecated term is an outdated or undesired term that in the past might be used as a synonym to a preferred term but now is not recommended (i.e. it is a not admitted non-preferred term). It is very important, however, to be careful with its handling, because complete removal would not allow access to historical material.

Note 3 to entry: Described in ISO 1087:2019 as deprecated term, term which is a synonym for a preferred term, but rated according to the acceptability rating scale as undesired.

3.20

document

any resource that can be classified or indexed in order that the data or information in it can be retrieved

Note 1 to entry: This definition refers not only to written and printed materials in paper or microform versions (for example, conventional books, journals, diagrams, maps), but also to non-printed media such as machine-readable and digitized records, Internet and intranet resources, films, sound recordings, records or collections of people, organizations, buildings, sites, monuments, three-dimensional objects as knowledge resources.

3.21

entry term

lead-in term

term ([3.77](#)) provided in a *controlled vocabulary* ([3.16](#)), not for direct use in metadata or *indexing* ([3.33](#)), but for the purpose of guiding the user to a *concept* ([3.14](#)), via another *term* ([3.77](#)) that can be used as a category label, subject heading, *concept label* ([3.15](#)), or *preferred term* ([3.56](#))

Note 1 to entry: Entry terms occurring in a thesaurus are generally known as non-preferred terms ([3.53](#)).

3.22

equivalence mapping

mapping that states that the *concept* ([3.14](#)) in the target vocabulary is considered identical in scope to the *concept* ([3.14](#)) in the source vocabulary.

cf. equivalence relationship ([3.26](#))

3.23

equivalence relationship

relationship between two *terms* ([3.77](#)) in a *thesaurus* ([3.78](#)) that both represent the same *concept* ([3.14](#))

Note 1 to entry: In ordinary discourse, terms that are quasi-synonyms may represent slightly different concepts. After inclusion in the thesaurus, however, the equivalence relationship clarifies that both are regarded as representing the same concept. When two or more such terms are in the same language within a monolingual or multilingual thesaurus, one of them is designated a preferred term and the other(s) non-preferred term(s); when two or more such terms are in the different languages of a multilingual thesaurus, each of them may be a preferred term in its own language respectively, and the relationship is known as cross-language equivalence.

Note 2 to entry: Entry terms are referred to as “alternative labels” in SKOS.

Note 3 to entry: Also described in ISO 1087:2019 – equivalence, relation between designations in different natural languages representing the same concept.

cf. equivalence mapping ([3.23](#))

3.24

exchange format

machine-readable format for representing information that is intended to facilitate exchange of the information between different applications

Note 1 to entry: The exchange format for a thesaurus often uses a markup language based on a standard such as XML (Extensible Markup Language) 0000, and is based on a data model for thesauri. While the data model provides a generic description of thesaurus structure and semantics, the exchange format expresses this in a formal language for the purpose of exchanging thesauri.

Note 2 to entry: Entry terms are referred to as “alternative labels” in SKOS and *equivalence mapping* ([3.25](#)).

3.25**facet**

grouping of *concepts* (3.14) of the same inherent category or “aspect-based classification”

cf. *node label* (3.49)

EXAMPLE

Animals, mice, daffodils and bacteria could all be members of a living organism’s facet.

Digging, writing and cooking could all be members of an actions facet.

Paris, the United Kingdom and the Alps could all be members of a places facet.

Note 1 to entry: Examples of high-level categories that can be used for grouping concepts into facets are: objects, materials, agents, actions, places and times.

3.26**facet analysis**

arrangement of subject areas into constituent *concepts* (3.14) grouped into *facets* (3.25), and the subdivision of *concepts* (3.14) into *narrower concepts* (3.47) by specified characteristics of division

3.27**facet indicator**

notational device that indicates the start of a new *facet* (3.25) or classification level, which functions as a node label and is not an index term. within a synthesized compound *notation* (3.51)

Note 1 to entry: Examples of facet indicators are the 0 in the Dewey Decimal Classification, and parentheses and quotation symbols in the Universal Decimal Classification. In the past, the term facet indicator has been used as synonymous with node label, but that usage is deprecated by ISO 25964, to avoid confusion

3.28**GUID (Globally Unique Identifier)**

16-byte string of characters that is generated in a manner that gives a high probability that the string is unique in any context

[SOURCE: ISO/IEC 2382-17:1999 replaced by ISO/IEC 2382:2015.]

Note 1 to entry: The term GUID is mostly used for Microsoft contexts. It is a 16-byte string GUID/UUID is a 128-byte string

3.29**hierarchical relationship**

relationship between a pair of *concepts* (3.14) of which one has a scope falling completely within the scope of the other, usually referred to as Broader and Narrower relationships in a thesaurus or taxonomy

Note 1 to entry: For example, in a hierarchical relationship "drug jars" is the broader context for "albarellos," which are a distinctive type of drug jar.

cf. *broader term* (3.4), *narrower term* (3.47)

Note 2 to entry: Several different types of hierarchical relationship exist. For a further explanation, see 11.2.

Note 3 to entry: Described in ISO 1087:2019 as “hierarchical relation, hierarchical concept relation (generic or partitive)”.

3.30**homograph**

one of a set of identical terms that refer to different entities

one of two or more words that are written in the same way, but have different meanings

EXAMPLE

In English: The word "bank" could refer to a financial institution or the side of a river.
In French: The word "avocat" could refer to a lawyer or to a fruit.
Lead a mineral or an element a lead for walking a dog, a management activity, a small body of water leading to another bigger one etc.)
Mercury (a God, a planet, an element, a type of car etc.),

Note 1 to entry: Homographs are sometimes referred to as homonyms, although the latter term applies more broadly, as it also includes pairs of terms such as "weights" and "waits" in English or "mer" and "mère" in French, which sound the same although they are spelt differently. C.f. ISO/IEC 2382-17:1999)

Note 2 to entry: homonym: term and definition standardized by ISO/IEC [ISO/IEC 2382-17:1999] .

3.31 identifier

set of symbols, usually numeric or alphanumeric, designating a *concept* (3.14) or a *term* (3.77) or another entity for purposes of unique identification within a determined context or resource, especially in a computer system or network

Note 1 to entry: A notation is sometimes used as an identifier.

Note 2 to entry: See 13.5 for further identifier descriptions.

3.32 index term

term (3.14) assigned to a document (3.23) in the process of *indexing* (3.33)

Note 1 to entry: Sometimes index terms are referred to as indexing terms, as keywords, descriptors or as tags, but the latter terms have other meanings too. Preferred terms from a thesaurus are very often used as index terms. It is also possible to use a code or a term Identifier.

Note 2 to entry: In SKOS they are used as concept labels.

3.33 indexing subject indexing

Analysis of the subject matter of a document (3.23) to identify the *concepts* (3.14) represented in it, and allocation of the corresponding *index terms* (3.32) to allow the information to be retrieved

Note 1 to entry: -The term "subject indexing" is often used for this concept, but as ISO 25964 does not deal with the indexing of other elements such as authors or dates, "indexing" is sufficient. Indexing can be carried out by human users or by automated agents.

3.34 information management

planning, collection, control, distribution and exploitation of information resources within an organization, including systems development, and disposal or long-term preservation

Note 1 to entry: See also ISO/IEC 2382:2015, definition 2121409.

3.35 information retrieval

the process of finding relevant information from a collection of data

action, methods and procedures for recovering information on a given matter or references to pertinent documents from stored data

c.f. 3.10.1.02 (ISO 5127) action, methods and procedures for recovering information on a given matter or references to pertinent documents from stored data

[SOURCE: ISO 8459:2009, definition 2.28]

3.36

information use

use of information can be defined as the interpretation and coding of environmental stimuli, whereby the human organism created new and adapted knowledge structures

(Cole 2008). September 13, 2010

3.37

interoperability

ability of two or more systems or components to exchange information and to use the information that has been exchanged

Note 1 to entry: Vocabularies can support interoperability by including relations to other vocabularies, by presenting data in standard formats and by using systems that support common computer protocols. See also ISO 25964-2.

3.38

lexical label

human readable label for a concept to provide easy recognition. See Clause 12.5.1 .

3.39

loan term

term ([3.14](#)) borrowed from another language that has become accepted in the borrowing language

EXAMPLE

"glasnost" is a Russian term that has become accepted in English
"gourmet" is a French term that has become accepted in English

A type of drinking cup, called in English "sturzbecher" is a loan term from German "Sturzbecher"

The term for the music genre originating in English as "jazz" is used as a loan term in Dutch, French, Italian, Spanish, and other languages

Kleenex a term borrowed from popular use meaning facial tissue

3.40

mapping

the process of connecting a concept from one source (usually the requestor) to a concept in another source (in this case the thesaurus preferred term).

Note 1 to entry: definition from ISO 25962 Part 2 clauses 3.40 and 3.41

3.42

mapping, gerund (verbal noun)

process of establishing relationships between the concepts (3.17) of one vocabulary and those of another

3.41

mapping, noun

(product of mapping process) relationship between a concept (3.17) in one vocabulary and one or more

concepts (3.17) in another

NOTE 1 A mapping generally has a direction, as discussed in Clause 6.

NOTE 2 Exceptionally, a mapping may involve a combination of two or more target vocabularies, where one or more of them act(s) as qualifier to the other (see 8.3.4).

3.41

markup

annotations or other type of encoding embedded in text, in conformity with a *markup language* (3.42)

3.42

markup language

set of encoding conventions that can be used to provide instructions for the interpretation of a text, by the use of annotations embedded in the text itself

Note 1 to entry: The interpretation often concerns issues such as content, structure or rendering of the text. Widely used examples include HTML (Hypertext Markup Language)⁰, which is largely concerned with presentation, and XML (Extensible Markup Language)⁰⁰⁰⁰, which addresses the structure of text.

3.43

metadata

data that identify attributes of a *document* (3.20) typically used to support functions such as location, discovery, documentation, evaluation and/or selection, information retrieval, management and use

Note 1 to entry: Preferred terms or notations selected during the indexing process are commonly applied as metadata values.

3.44

mono-hierarchical structure

hierarchical arrangement of *concepts* (3.14), in a thesaurus (3.77) or classification scheme (3.6), in which each *concept* (3.14) can have only one broader concept (3.4) at the level immediately above

cf. *polyhierarchical structure* (3.53)

EXAMPLE In a mono-hierarchical structure, the concept of pianos cannot be narrower to keyboard instruments as well as narrower to stringed instruments; a choice has to be made of one of these concepts to determine its single relationship.

3.45

multilingual thesaurus

thesaurus (3.78) in which *terms* (3.77) concepts and terms with their relational structures are available in two or more natural languages

3.46

multi-word term

term (3.77) consisting of more than one word separated by spaces.

cf. *compound term* (3.13)

Note 1 to entry: For information retrieval these words need to be searched as a phrase to maintain meaning not as single words later combined with a Boolean AND combine operations.

EXAMPLE

cost benefit analysis French Renaissance Revival monumentale beeldhouwwerken [Dutch] campanella celtica [Italian] brise-soleils [French]
--

Note 2 to entry: Compare with ISO 1087:2019 – multi-word term, complex term that consists of more than one word.

3.47**narrower concept**

concept ([3.14](#)) that is narrower than the one in question

Note 1 to entry: The scope of the narrower concept falls completely within the scope of the broader.

Note 2 to entry: Described in ISO 1087:2019 as “subordinate concept, narrower concept, specific concept or partitive concept.”

3.48**narrower term**

preferred term ([3.56](#)) representing a concept ([3.6](#)) that is narrower than the one in question

Note 1 to entry: The hierarchical relationship should always be reciprocal, narrower term and Broader ([3.4](#)) need to both be listed in the thesaurus. The relationship between the two is commonly indicated with the tag NT. For more explanation see [11.2.1](#)

EXAMPLE

in the example, each child must be completely contained in the meaning and scope of the parent and also the grandparent concept (e.g., "place brick" is a type of "soft brick," and also a type of "fired brick")

```

fired brick
...hard-burned brick
.....clinker brick
...soft brick
.....grizzle
.....place brick
.....sammel bric
  
```

Fruit

```

apple
pear
  
```

3.49**node label**

label inserted into a hierarchical or classified display to show how the *terms* ([3.77](#)) have been arranged

Note 1 to entry: A node label is neither a preferred term nor a non-preferred term. It is a placeholder term not used for indexing which provides separation and maintains the hierarchy. It contains one of two different types of information:

- the name of a facet to which following terms belong; or
- the attribute or characteristic of division by which an array of sibling concepts has been sorted or grouped.

See examples in [Clause 13](#).

3.50

non-preferred term

term (3.77) that is not assigned to *documents* (3.20) but is provided as an entry point in a *thesaurus* (3.78) or index as a term with an equivalent meaning to the preferred to which it directs to and to which it is related with the *equivalence relationship* (3.23). Usually a synonym but could be an outdated and now deprecated term. Sometimes referred to as a “non-descriptor.”

cf. *entry term* (3.21) cf. preferred term (3.58) for parallel cf. note at 3.56 preferred term: cf. *non-preferred term* (3.50)

EXAMPLE

hounds
USE dogs
<i>illegal aliens</i>
USE undocumented immigrants

Note 1 to entry: In this example, "hounds" is a nonpreferred term, while "dogs" is the preferred term (or concept label) that should be used in its place. See also [Section 8.3.1](#).

Note 2 to entry: Compare with ISO 1087:2019 – admitted term, term that is a synonym for a preferred term, but not rated according to the acceptability rating scale as a preferred term.

3.51

notation

Also known as notation, classmark class code, class number

set of symbols representing a *concept* (3.14) in a *structured vocabulary* (3.71), especially a *classification scheme* (3.10)

EXAMPLE

Notation	Source vocabulary	Concept
07.04.4	ILO Thesaurus	fishery policy and development
622.342 2	Dewey Decimal Classification	gold mining
373.3.016:51	Universal Decimal Classification	mathematics curriculum in primary schools
SBS XEJ B	Bliss Bibliographic Classification	endangered species law
9C61 H40-H42 USE ICD 11	International Statistical Classification of Diseases and Related Health Problems	glaucoma

Note 1 to entry: Notation is sometimes used to sort and/or locate concepts in a predetermined systematic order and, optionally, to display how the components of complex concepts have been structured and grouped. A notation can provide the link between alphabetical and systematic lists in a thesaurus. In the context of classification schemes, "concepts" are often known as "subjects", especially when they are complex, as in the examples above.

3.52

paradigmatic relationship

relationship between *concepts* (3.14) that is inherent in the *concepts* (3.14) themselves whether by an explicit associative or hierarchical relationship or the potential to be so related in a thesaurus. Also known as “a priori relationship”

EXAMPLE (with the associative relationship)

a "lithographer" creates "lithographs" using the process of "lithography"

Need another example which uses different terms

Note 1 to entry: Such relationships are shown in a structured vocabulary, independently of any indexed document. For a more complete discussion of paradigmatic and syntagmatic relationships, see [5.3](#).

3.53

polyhierarchical structure

hierarchical arrangement of *concepts* ([3.14](#)), in a *thesaurus* ([3.78](#)) or classification in which each concept ([3.16](#)) can have more than one *broader concept* ([3.3](#))

cf. *monohierarchical structure* ([3.44](#))

EXAMPLE

In a polyhierarchical structure, organs (musical instruments) could be listed under keyboard instruments as well as under wind instruments.

Wind instruments

— **Organs**

Keyboard instruments

— **Organs**

In a polyhierarchy, the African ceremonial swords of the Asante people is a narrower context of both "ceremonial weapons" and "edged weapons"

Note 1 to entry: In a polyhierarchical structure, a single concept can occur in more than one place in the hierarchical structure of the thesaurus. Its attributes and relationships, and specifically its narrower and related terms, are the same wherever it occurs.

3.54

post-coordination

combination of *preferred terms* ([3.56](#)) of a *controlled vocabulary* ([3.16](#)) at the time of searching which reflect their combination of application to the same documents when they were indexed, but in an unstructured manner

cf. *pre-coordination* ([3.55](#))

EXAMPLE

The post-coordinated search expression "microwaves AND radiation" can be used to retrieve documents on microwave radiation, when these have been indexed under the separate terms "microwaves" and "radiation" rather than a compound term.

A search on "blue silk kimono" combines: color + fabric + clothing item

3.55

pre-coordination

combination of *concepts* ([3.14](#)), classes or *terms* ([3.77](#)) of a *controlled vocabulary* ([3.16](#)) at the time of its construction or at the time of using it for indexing ([3.36](#)) or *classification* ([3.8](#)) in a structured manner *which standardises the meaning given by the combination* where a term is associated with another term, such as a subentry or subdivision of the other

cf. *post-coordination* ([3.54](#))

EXAMPLE 1

The class "general theory", when placed within the broader class "music", refers only to the pre-coordinated subject "theory of music" and not to theory in general.

EXAMPLE 2

The pre-coordinated string "cardboard – recycling" might appear in a subject heading scheme or, if not enumerated there, might be synthesized by an indexer when needed for a particular document.

Examples can be found in the Library of Congress Subject Headings and BLISS, and back of the book indexing.

3.56**preferred term**

term (3.77) used to represent a *concept* (3.14) when *indexing* (3.33) and the term that is displayed in hierarchical displays of a thesaurus or taxonomy, which do not support the display of more than one term for a concept

cf. *non-preferred term* (3.50)

Note 1 to entry: A preferred term is a noun or noun phrase. See also *Concept label* (3.15) and [Section 7.3.1](#).

Note 2 to entry: Described in ISO 1087:2019 as “preferred term, term rated according to the acceptability rating as the primary term for a given concept.”

Note 3 to entry: Also known as a descriptor.

3.57**protocol**

convention that defines the syntax, semantics and synchronization of the communication process between two computers in order to enable a particular service

EXAMPLE

An Application Protocol (Programming) Interface or API is a frequent handshake between two computers.

3.58**quasi-synonym**

one of two or more *terms* (3.77) in a given controlled vocabulary (3.19) whose meanings are generally regarded as different in ordinary usage, but which may be treated as designators for the same *concept* (3.14), are regarded as almost identical in a given controlled vocabulary (3.19)

Note 1 to entry: May also be called a near-synonym.

EXAMPLE

diseases, disorders earthquakes, earth tremors

3.59**related concept**

concept (3.14) that has an *associative relationship* (3.2) than the one in question.

EXAMPLE

For transitional phases between Stone Age and Bronze Age, an associative relationship exists where "Chalcolithic" is overlapping in scope with "Early Bronze Age".

Note 1 to entry: this is also the name used in SKOS

3.60**related term**

preferred term (3.56) representing a *concept* (3.14) that has an *associative relationship* (3.2) with the one in question

Note 1 to entry: The relationship between related terms is commonly indicated with the tag RT. For a further explanation, see 11.3.

3.61**relationship**

relation

intellectual connection between two or more *concepts* (3.14) or between two or more terms representing the same concept

Note 1 to entry: Described in ISO 5127:2017 where the relationship is in general defined as the connection between two or more elements. In this standard the connection between concepts (hierarchical, associative) and between terms designation one particular concept (equivalence etc) are examined.

3.62**schedule**

terms (3.77), *notations* (3.51), captions, cross-references and *scope notes* (3.63) set out to exhibit the content and structure of a *structured vocabulary* (3.71)

3.63**scope note**

note that defines or clarifies the semantic boundaries of a *concept* (3.14) as it is used in the *structured vocabulary* (3.71)

EXAMPLE

For "landscape architects": People who specialize in designing the scenic environment on a relatively large scale, including the layout of plants and accompanying structures. Distinguished from "garden designers."

Note 1 to entry: A term used to label a concept can have several meanings in normal usage. A scope note is used to restrict the concept to only one of those meanings, and where necessary refers to other concepts that are included or excluded from the scope of the concept being clarified. Automated applications may be assisted by Scope notes which give (in a consistent manner) examples of sentences with the different meanings and using disambiguating words/qualifiers. With automation including artificial intelligence applications the problem differs. When a concept which can have several meanings is resolved by the algorithms of co-occurrence in automation using the surrounding context. The use of disambiguating words/qualifiers and with synonyms from the thesaurus and of controlled vocabularies in general helps this automation be more specific and less prone to hallucinations. In Scope notes we define the meaning of terms and their different senses. A natural language example is helpful and provides content of the meaning of that particular use in the instance of a particular term. Rather than using parenthesis for guidance, a thesaurus with examples of word uses in an example. Parenthetical qualifiers are not recommended in thesaurus use since they are not useful in computer search retrieval systems.

3.64**search term**

term (3.77) forming all or part of a search query as a string of characters entered by a searcher. The computer will usually parse "words" as strings of characters separated by a space.

Note 1 to entry: c.f.12.4.4 Normalization for information retrieval matching a string of characters entered by a searcher (search term) with a string of characters extracted from metadata and stored into an index (index term). The search term and the index term sometimes differ.

Note 2 to entry: Existing vocabulary resources should be researched before construction begins.

Note 3 to entry: The transaction log of a relevant website, showing the search terms most frequently employed by users.

3.65

sequence number

automatically generated numerical identifiers unique to a single structured vocabulary. typically for each term but in a concept-based system, could be for each concept

See clause 12.5.2

3.66

sibling concept

one of two or more *concepts* (3.14) with the same immediate *broader concept* (3.3), each of these being represented by a *preferred term* (3.56)

EXAMPLE

In the example, “ballets” and “minuets” are siblings of each other, and “butterfly dances” and “hoop dances” are siblings as they are at the same levels of the hierarchy and both narrower terms to the same broader term.

```

_dances
...ballets
...bugaku
...classical dances
...equestrian ballets
...minuets
...Native American dances
.....butterfly dances
.....ghost dances
.....hoop dances
.....stomp dances
    
```

Note 1 to entry: Compare with ISO 1087:2019 – coordinate concept, subordinate concept resulting from the same criterion of subdivision as another subordinate concept

3.67

sibling term

one of two or more *preferred terms* (3.56) with the same immediate *broader term* (3.4)

EXAMPLE

In the following, chairs and tables are sibling terms in the same array, while no siblings are shown for “furniture”, “armchairs” or “dining tables”: That is they occupy the same level in the hierarchy

```

furniture
  chairs
    armchairs
  tables
    dining tables
    
```

3.68

SKOS

used for thesauri, classification schemes, taxonomies, and other vocabulary types

Note 1 to entry: SKOS is based on the RDF (Resource Description Framework) data model, and may be encoded in any syntax supported by RDF, including XML, JSON, Turtle, etc.

Note 2 to entry: more information in clause 15.3.8 exchange formats.

3.69

source language

language serving as a starting point in translation or in a search for *term* (3.77) equivalents.

c.f. Contrast with *target language* (3.75)

3.70

specificity

capability of a *structured vocabulary* (3.71) to express a subject in depth and in detail, also referred to as granularity

EXAMPLE

in sewing terminology "chain stitches," "couched stitches," and "French knots" are more specific concepts than "stitches"

Note 1 to entry: For a further explanation, see the discussion of specificity in 8.1 and other places.

3.71

structured vocabulary

organized set of *terms* (3.77), headings or codes representing *concepts* (3.14) and their inter-relationships, which can be used to support *information retrieval* (3.35), *information management* (3.34), and *information use* (3.36)

Note 1 to entry: A structured vocabulary can also be used for other purposes. In the context of information retrieval, management, and use the vocabulary needs to be accompanied by rules for how to apply the terms. Various types of structured vocabulary are addressed in ISO 25964-2, including classification schemes, subject heading schemes, etc.

3.72

subject heading scheme

structured vocabulary (3.71) comprising *terms* (3.77) available for subject *indexing* (3.33), plus rules for combining them into pre-coordinated strings of *terms* (3.77) where necessary. Also referred to as subject heading language, subject heading list, SHL

Note 1 to entry: For example, "Laws--Portugal--eighteenth century" is a subject heading comprising terms for a topic, the location, and the time span.

Note 2 to entry: Subject Headings are discussed in depth in ISO 25964 Part 2

3.73

synonym

one of two or more *terms* (3.77) denoting the same *concept* (3.14)

EXAMPLE

In English: guarantees, warranties heart attack, myocardial infarction HIV, human immunodeficiency virus Stage IV breast cancer, Metastatic breast cancer, Invasive breast cancer
In French: schiste, phyllade VIH, virus de l'immunodéficience humaine crise cardiaque, infarctus du myocarde In German: "Viola," "Altgeige," "Armgeige," and "Bratsche"

Note 1 to entry: Abbreviations and their full forms can be treated as synonyms. See also [Section 7.3](#).

Note 2 to entry: Also described in ISO 1087:2019 – synonymy, relation between designations in a given natural language representing the same concept.

3.74

syntagmatic relationship

relationship between *concepts* ([3.14](#)) that exists only because they occur together in a *document* ([3.20](#)) being indexed. Also known as posteriori relationship

Note 1 to entry: Such relationships are not generally valid in contexts other than the document being indexed, and therefore they do not form part of the structure of a thesaurus. For a more complete discussion of syntagmatic and paradigmatic relationships, see 0.

EXAMPLE The words "gem-encrusted reliquary" may be used to describe a given object, but the condition of "gem-encrusted" is not necessary or typical for the concept of "reliquary".

3.75

target language

language providing a translation or an equivalent for a *term* ([3.77](#)) existing in a *source language* ([3.69](#))

3.76

taxonomy

collection of controlled vocabulary terms organized into a hierarchical structure

Note 1 to entry: A taxonomy is also a structured vocabulary. The primary purpose of a taxonomy is to provide a navigational structure for the user, to guide searchers to the desired concept

Note 2 to entry: Each term in a taxonomy is in one or more parent/child (broader/narrower) relationships to other terms in the taxonomy.

[SOURCE: ANSI/NISO Z39.19-2005 r2010 Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies, [section 4.1](#) Definitions]

3.77

term

word or phrase used to represent / designate a *concept* ([3.14](#))

EXAMPLE

schools school uniform costs of schooling teaching

Note 1 to entry: Thesaurus terms can be either preferred terms or non-preferred terms.

Note 2 to entry: Compare with ISO 1087:2019 – term, designation that represents a general concept by linguistic means.

3.78

thesaurus

controlled ([3.16](#)) and *structured vocabulary* ([3.71](#)) in which *concepts* ([3.14](#)) are represented by *terms* ([3.77](#)), organized so that relationships between *concepts* ([3.14](#)) are made explicit, and *preferred terms* ([3.56](#)) are accompanied by lead-in entries for *synonyms* ([3.73](#)) or *quasi-synonyms* ([3.58](#))

Note 1 to entry: The purpose of a thesaurus is to guide both the indexer and the searcher to select the same preferred term or combination of preferred terms to represent a given concept. For this reason, a thesaurus is optimized for human navigability and terminological coverage of a domain.

3.79

top term

preferred term ([3.68](#)) representing a *concept* ([3.14](#)) that has no *broader concept* ([3.3](#)) in the *thesaurus* ([3.78](#))

EXAMPLE

“Objects Facet” is the top of the hierarchy

Objects Facet

...Built Environment

...Components

...Furnishings and Equipment

...Object Genres

...Object Groupings

.....archives

.....assemblages

.....collections

.....art collections

.....bookstocks

.....fonds

3.80

UUIDs

Universally Unique Identifiers

identifier that is unique with respect to the space of all universal unique identifiers. See clause 12.5.4

[SOURCE: ISO/IEC [ISO/IEC 2382-17:1999] replaced by ISO/IEC 2382:2015.]

3.81**vocabulary control**

management of a vocabulary in order to disambiguate and constrain the form of the *terms* (3.77) and limit the number of *concepts* (3.14) and *terms* (3.77) available for *indexing* (3.33) *without compromising indexing quality*

Note 1 to entry: Control is achieved by distinguishing between homographs, so that each one has just one meaning, and by picking out from a set of synonyms or quasi-synonyms, the one which is to be preferred for use in indexing. The purpose of these restrictions is to increase the likelihood of indexers and searchers choosing the same term to label a particular concept.

3.82**warrant**

choice of the *concept label* (3.15) to be used based on the needs of the particular subject domain and community of users

Note 1 to entry: “Literary warrant” is the term (3.78) that most commonly appears in publications. “Organizational warrant” is the official term (3.78) chosen based on organizational needs, priorities or policies. “User warrant” is the term (3.78) users most commonly use.

4 Symbols, abbreviated terms and other conventions

4.1 Table 1 shows the abbreviated terms that are used in English language examples of thesaurus records as tags to prefix terms and notes. Each tag indicates the relationship or function of the term or note that follows.

Table 1 — Symbols and abbreviations

Description	Symbol	Tag	Meaning
<i>Descriptive elements</i>		SN	Scope note
		DEF	Definition
		HN	History note
<i>Codes</i>		SC	Subject category; a code or notation applied to a group of concepts relevant to a particular subject
		CC	Concept code or notation
	→	PT	Use; the term that follows the tag is the preferred term that should be used in place of the non-preferred term that precedes the tag
	=	NPT	Use for or Used for; the term that follows the tag is a non-preferred term for which the preferred term preceding the tag should be used instead
<i>Relationships</i>		USE...+	The two or more preferred terms following the tags should be used together to represent the indicated concept
		UF+	The non-preferred term that follows should be represented by a combination of preferred terms including the preferred term that precedes the tag
		TT	Top term; the preferred term that follows the tag represents the broadest concept in a hierarchy to which the specific concept belongs
	<	BT	Broader term; the term that follows the tag represents a concept having a wider meaning
		BTG	Broader term (generic)
		BTI	Broader term (instantial)
	→<	BTP	Broader term (partitive)
	>	NT	Narrower term; the term that follows the tag refers to a concept with a more specific meaning
		NTG	Narrower term (generic)

Table 1 (*continued*)

Description	Symbol	Tag	Meaning
		NTI	Narrower term (instantial)
	>–	NTP	Narrower term (partitive)
	—	RT	Related term; the term that follows the tag is associated, but is not a synonym, a quasi-synonym, a broader term or a narrower term

Table 2 — English language tags and their equivalents in other languages

Tags in English	Tags in French	Tags in German	Tags in Danish	Tags in Finnish	Tags in Norwegian	Tags in Swedish	Tags in Spanish	Tags in Chinese	Tags in Russian	Tags in Arabic
SN	NE	H	SN		NO	SN	NA	J	лп	ت
Scope note	Note explicative ^b <i>Note d'emploi</i>	Erläuterung (Hinweis)	Forklarende note	Huomautus ^c	Note	Anmärkning	Nota de aplicación <i>Nota de alcance</i>	解jie	Лексическое примечание	تبصرة
PT	EM	BS	USE	KÄ	BRUK	USE	USE	Y	см	س
Use	Employer	Benutze Gattungsbegriff	Brug	käytä	Bruk	Använd	Use	用yong	Смотри	استخدم
UF	EP	BF	UF	KT	BF	UF	UP	D	С	ل م
Use for	Employé pour <i>Employer pour</i>	Benutzt für <i>Benutze für</i>	Brugt for	korvattu termi	Brukt for	Används istället för	Usado por <i>Use por</i>	代dai	Синоним	مستخدم ل
BT	TG	OB	BT	LT	OT	BT	TG	S	В	ع م
Broader term	Terme générique	Oberbegriff	Overordnet term	laajempi termi	Overordnet term	Överordnad term	Término genérico	属shu	Выше	مصطلح اعم
NT	TS	UB	NT	ST	UT	NT	TE	F	Н	م ض
Narrower term	Terme spécifique	Unterbegriff Individualbegriff	Underordnet term	suppeampi termi	Underordnet term	Underordnad term	Término específico	分fen	Ниже	مصطلح اضيق
RT	TA	VB	RT	RT	SO	RT	TR	C	a	
Related term	Terme associé	Verwandter Begriff Bevorzugter Begriff	Relateret term	rinnakaistermi	Se også	Relaterad term	Término relacionado	参can	Ассоциация	
BTG	TGG	OA	BTG						вр	
Broader term (generic)	Terme générique (générique)	Oberbegriff (Abstraktionsrelation) Definition	Overordnet term (generisk)						Выше—род	
BTP	TGP	SP	BTP						вц	

NOTE The selection of languages in Table 2 is open-ended. Future editions of this part of ISO 25964 could include additional tags.

^a Lines in italics show variant tags and/or expansions.^b French AFNOR Standard (AFNOR Z47-100) proposed NA: Note d'application^c In the Finnish standard "Scope note" is translated "Selitys" at the moment, but "Huomautus" would be a better translation. There is no associated tag.https://agrovoc.fao.org/browse/agrovoc/en/page/2uri=http%3A%2F%2Faim.s.fao.org%2Ffaos%2Fagrovoc%2Fc_c37670aa<https://vocabularies.unesco.org/browser/thesaurus/fr/page/2uri=http%3A%2F%2Fvocabularies.unesco.org%2Fthesaurus%2Fconcept899><https://vocabularies.unesco.org/browser/thesaurus/en/page/2uri=http%3A%2F%2Fvocabularies.unesco.org%2Fthesaurus%2Fconcept7343>

Full document details: 25/30472169 DC

AVA Documents
Document Sharing Hub

4.2 Thesauri in Dutch use the same tags as in English, without translation.

4.3 The abbreviations listed as tags in [Table 2](#) have acquired status as generally recognized conventions, and they occur in many published thesauri. They have mnemonic value but apply only in particular languages. If a more neutral system of abbreviation is required, the language-independent symbols in the left-hand column of [Table 1](#) are recommended.

4.4 The following conventions are also used in examples throughout this part of ISO 25964. Note that these conventions are not required of all thesauri.

Preferred terms are printed in the case they would normally have in full text. For most languages, that means lower case throughout, with the exception of upper case for abbreviations and acronyms, and initial capital letters for proper names.

EXAMPLES

cars
animals

Non-preferred terms are printed in *italics*.

EXAMPLES

cars	<i>automobiles</i>
UF <i>automobiles</i>	USE cars
animals	<i>fauna</i>
UF <i>fauna</i>	USE animals
World Health Organization	<i>WHO</i>
UF <i>WHO</i>	USE World Health Organization
German, Ale-Gläser	<i>Biergläser</i>
BF <i>Biergläser</i>	BS Ale-Gläser

5 Thesaurus overview and objectives

5.1 Overall objective

The aim of a thesaurus is to describe content with the term or terms which represent the same concepts. As well as to guide the indexer and the searcher. To achieve this, a thesaurus gathers and arranges all the concepts that might be useful in a given domain. The concepts are represented by terms, organized so that relationships between concepts are made explicit (see [11 Relationships between concepts](#)). For each concept, one of the terms is selected as the preferred term which functions as the concept label (see [7.6](#)).

When implemented in a system, the terms and relationships in the thesaurus may be exploited in numerous ways, including:

- concept expansion or refinement;
- suggestion of alternative terms;
- support for clustering or filtering or other means of refocusing a query;
- identification of common spelling mistakes;
- support of automatic indexing and categorization.

The interplay between terms, concepts and relationships is illustrated in the data model in [Clause 15](#). For reasons of practicality this part of ISO 25964 refers frequently to "terms". But it should never be forgotten

that the purpose of manipulating terms is to support information retrieval, terminology management, categorization, and semantic integration using the underlying concepts.

5.2 Vocabulary control and its purpose

Establishing an appropriate preferred term to represent a particular concept is not always straightforward because a concept can often be expressed in more than one way. Furthermore, in ordinary discourse one term can have more than one meaning. Vocabulary control is therefore essential, and thesauri are used to achieve this by the following three principal means.

1. Concepts and terms are deliberately restricted in scope to selected meanings. Unlike the terms in a dictionary, which are often accompanied by a number of different definitions reflecting common usage, each term in a thesaurus is generally restricted to whichever single meaning serves the needs of a retrieval, management, and use system most effectively. The structure of a thesaurus, notably its display of hierarchical relationships, frequently indicates the intended meaning of a term.
2. The synonym ring or assembled variant terms also indicates the intended meaning of a term. If this technique is not sufficiently explicit, a scope note should accompany the term. This note should state the chosen meaning, and it may also indicate other meanings that are recognized in natural language, but which have been deliberately excluded for indexing purposes.
3. When the same concept can be expressed by two or more synonyms or quasi-synonyms in the same language, one of these terms is usually selected as the preferred term, which is then used consistently in indexing or simply as the first or only term representing the concept. Reference to the preferred term should be made from any synonym that might also function as a user's access point.

The thesaurus must reflect the user's content and perspective. To do otherwise will mean that the work is not used by the community it was built for. Appropriate terms will vary by audience and content. Thesauri should be built for use not to dictate how the users speak and ask them to change their ways. The thesaurus has an important role in mediating between terms used in discourse and those that function effectively for information retrieval, management, and use. To achieve the retrieval, management, and use benefits, users need to accept a degree of artificiality in the controlled vocabulary (although in some systems this difficulty can be overcome by automatic substitution of the user's term with the preferred form).

So that a thesaurus can function effectively in a multilingual context, the concepts included need to be represented in all of the languages present, enabling speakers of these languages to have access to them. If the thesaurus is "symmetrical" as described in this part of ISO 25964 and modelled in Clause 16, each concept should have a preferred term in each language, and the scope should be the same in all the languages. This constraint sometimes adds to the artificiality of the vocabulary. Alternative approaches (for example the use of mappings between vocabularies) are described in ISO 25964-2.

5.3 Paradigmatic versus syntagmatic relationships

The purpose of establishing and displaying relationships is likewise to guide users (or intelligent agents acting on their behalf) to choose the most appropriate term(s) for expressing a given concept.

This is achieved by presenting terms that a user might consider instead of, or as well as, the term entered first. Means for displaying relationships are described in [Clause 14](#).

Two kinds of relationship are commonly encountered between indexing terms in any one language.

1. Relationships that exist only because the concepts occur together in the context of a particular document are called syntagmatic relationships. When two or more indexing terms are assigned to the same document, this act effectively records a syntagmatic relationship between them. In a thesaurus, however, inclusion of syntagmatic relationships is not recommended.

EXAMPLE 1

An indexer dealing with a work on “Computers in banks in Amsterdam” might identify three concepts represented respectively by the terms: “banks (financial institutions)”, “computers” and “Amsterdam”. He assigns these to the document. In a post-coordinate system, the relationship between these terms is not explicitly indicated in metadata, but the document would nonetheless be retrieved if any or all of these terms were used as retrieval keys. In a pre-coordinated index, an entry can be provided for the combination of three terms, with a reference to the location where the corresponding document can be found. The concepts in this example are not normally associated according to common frames of reference, and their interrelationships can, therefore, be regarded as document dependent.

- Relationships that are valid in almost all contexts, especially when they are inherent in the concepts that the terms represent, are known as paradigmatic relationships. It is useful to show paradigmatic relationships between thesaurus concepts, as they will frequently guide users to concepts closely related to the terms they first thought of.

EXAMPLE 2

Using concepts from the previous example, “banks (financial institutions)” has an inherent relationship with the broader concept “financial institutions”; “computers” has a strong association with “data processing”, independent of the subject of the document being indexed; and “Amsterdam” is inseparably linked with “Netherlands”. A user looking up any of these linked terms might be interested in information indexed with the former terms. These paradigmatic relationships are independent of any particular document. They are generally recognized and could be established through reference to standard works, such as dictionaries and encyclopedias.

The distinction between these two kinds of relationships can be displayed as shown In [Figure 1](#).

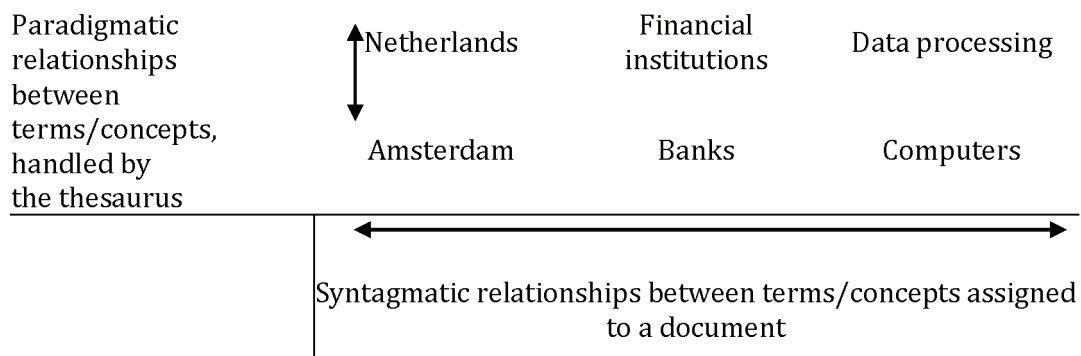


Figure 1 — Paradigmatic and syntagmatic relationships

5.4 Types of paradigmatic relationship

Three types of paradigmatic relationship between terms, or between concepts, should be established and clearly distinguished:

- the equivalence relationship as it applies in both monolingual and multilingual situations (see [Clauses 9](#) and [10](#) respectively);
- the hierarchical relationship (see [11.2](#));
- the associative relationship (see [11.3](#)).

Further subdivisions of each of these classes are described in [Clauses 9](#) to [11](#). Each of these relationships should be established reciprocally, and this should be indicated by a system of tags, symbols or abbreviations

used to represent relationships in a thesaurus. The conventional tags described in [Clause 4](#) are used in the examples that follow.

6 Concepts and their scope in a thesaurus

6.1 Conceptual basis

6.1.1 The prime application for a thesaurus is in information retrieval, management, and use, where the aim is to search for concepts. As explained in [5.1](#) and [5.2](#), the concepts are represented by terms. Each term included in a thesaurus should represent a single concept (or unit of thought). Concepts can range from the simple (e.g. "cats") to the very complex (e.g. "racial discrimination among ethnic minorities"). Compound terms or phrases are generally needed to express the more complex concepts. See [Clause 11](#), Relationships between concepts, for a more complete discussion of these.

6.1.2 The concepts included in a thesaurus may be considered as belonging to mutually exclusive categories based on shared characteristics. The following examples are indicative but do not list all the possible categories.

1. Things and their physical parts.

EXAMPLE SET 1

birds
birth certificates
limbs
microforms
monuments
mountain regions

NOTE Physical parts are also things and can have their own parts.

Materials.

EXAMPLE SET 2

adhesives
rubber
titanium

Activities or processes.

EXAMPLE SET 3

dressmaking
fertilization
glaciation
land management

Events or occurrences.

EXAMPLE SET 4

birthdays
civil wars
revolutions

Properties of persons, things, materials or actions.

EXAMPLE SET 5

consciousness
elasticity
personality
speed

Disciplines or subject fields.

EXAMPLE SET 6

archaeology
organic chemistry
theology

Units of measurement.

EXAMPLE SET 7

hertz
kilometers

Types of people and organizations.

EXAMPLE SET 8

charities
children
international nongovernmental organizations
nations
poets
visually impaired people

Unit of Time

EXAMPLE SET 9

hours
minutes
nanoseconds

6.1.3 Unique entities expressed as proper names may also be included in a thesaurus, where they should be used only for records that are *about* the person, document or object, etc. When the entity represented by the proper name is not the subject of a document, but has a different relationship, such as being its creator or title, the proper name should not be recorded together with subject terms, but elsewhere.

Proper names also belong to categories such as those listed in [6.1.2](#).

1. Places.

EXAMPLE SET 1

Australia Milky Way South Kensington Sri Lanka

2. Specific objects, topographical features, and other entities.

EXAMPLE SET 2

Magna Carta Mona Lisa Nelson's Column Romeo and Juliet Skylab

3. Individuals, organizational posts, and corporate bodies.

EXAMPLE SET 3

Burns, Robert United Nations Secretary-General World Health Organization
--

4. Events, holidays, and conferences.

EXAMPLE SET 4

Vietnam War Night of Mawlid Halloween Paris Climate Change Conference
--

6.2 Scope notes

The intended scope of a concept in the thesaurus is not always the most common meaning associated with its preferred term (or any one of its non-preferred terms) in ordinary discourse. In the thesaurus, context is usually provided by the hierarchy of broader and narrower concepts linked to the concept in question and this helps to elucidate the intended scope. Often it is possible to formulate the preferred term of the concept in an unambiguous way that conveys the intended scope to any user. However, in cases where these measures are not appropriate or sufficient, or where additional information would help to clarify the meaning and make the usage more consistent, an explicit scope note should be used.

A scope note should be used to clarify the boundaries of a concept, especially when the meaning of the preferred term in ordinary discourse can be interpreted too broadly or too narrowly, or to distinguish between preferred terms that have overlapping meanings in natural language. It can also be used to provide other advice on term usage to either the indexer or the searcher. In order to assist automated indexing and AI applications, it may be helpful to provide natural language examples of the terms in sentences for the different meanings, as part of the scope note. A scope note need not be a full definition but should clarify the

intended use of a term within the thesaurus. Other types of notes may be conveyed separately, for definitions and for history notes.

Unlike the qualifiers considered in a scope note is not regarded as forming part of the term to which it is attached.

EXAMPLE 1

microwave frequencies SN 1 GHz to 300 GHz

EXAMPLE 2

illuminations SN Includes both the ornamental decoration and the illustrations in manuscripts, as well as in some early printed books, if done by hand
--

6.3 Reciprocal scope notes

When reference is made to other concepts in a scope note, a reciprocal scope note should generally be provided for each concept mentioned.

EXAMPLE

In English: food products SN Use only for products intended for human consumption. For products for animals, see petfoods or feeds . petfoods SN Food products for animals maintained as domestic pets. For products for human consumption, see food products . For products intended for non-domestic animals, see feeds . feeds SN Products intended for non-domestic animals. For products for domestic pets, see pet foods . For products for human consumption, see food products .
In French: statistique NE Désigne la science statistique; pour les données statistiques employer “statistiques” statistiques NE Désigne les données statistiques; pour la science statistique, employer “statistique”

Even where the scope of only one of the concepts requires clarification, it is useful to make an editorial note in the record for any additional concept, of where it has been cited. The purpose of the reciprocal reference is to ensure that when a change is made to one of the concepts, or it is deleted, the effect on the other concept is considered. Although in general, it can be based on the common use. General thesauri serve different communities and languages, and ultimately a decision must be taken on which use will prevail for the concept label for the community the thesaurus serves.

7 Thesaurus terms

7.1 Form of terms

The terms selected to represent concepts may be single-word terms or multi-word terms (see [Clause 8](#)).

7.2 Term attributes

Terms may have a set of optional attributes conveying information on the source or nature of the term. In displays (or user interfaces) for a thesaurus, these attributes may be represented as flags, sometimes represented as capital letters or symbols in parentheses following the term. Some attributes intended solely for thesaurus administration purposes may not be displayed to end-users. Normally the thesaurus administrator specifies the type of attribute from a controlled list. Employing a range of term attributes can facilitate a wide entry vocabulary which will assist query expansion in search and auto-suggestion or named entity recognition in machine aided indexing applications.

7.3 Selection of the preferred term for display as concept label

7.3.1 General

Usually more than one term in the same language is available to represent the same concept. The situation across languages is described in [Clause 10](#). When multiple terms are included corresponding to the same concept, one of these should be designated the default Preferred term, in each language), for display as a Concept label. Where there is a choice between alternative terms, the preferences of the community to be served by the thesaurus should be adopted (subject to the avoidance of ambiguity or language that might be offensive to some users). Ascertaining such preferences can require some consideration and time (see [Section 14.3.5](#) on DEI considerations).

If possible, the choice of the preferred term to act as the concept label should be configurable by the administrator or user by changing a concept label flag when setting up the thesaurus or adapting the user interface. In alphabetical displays, a reference can be made from other terms to the preferred term. Alternative terms are sometimes referred to as Entry Terms. The term attribute types for the thesaurus can assist a consistent choice of preferred terms across concepts.

Attribute types may include plural, singular, other forms of speech, variant spellings, common spelling mistakes, synonyms of different etymological roots, scientific, vernacular and sometimes antonyms. For example, museum cataloguing traditionally uses the singular and libraries the plural form of the term when indexing. Indexing applications should offer the capability of indexing with a term identifier. Commonly encountered term attribute types are outlined immediately below and may form the basis of a controlled list appropriate to the users of the thesaurus (and see also discussion on equivalence in [Section 9](#)).

This will not be true for all languages and subject areas.

EXAMPLE 1

museum cataloguing traditionally uses the singular and libraries the plural form of the term when indexing.

EXAMPLE 2

In Germany in libraries the singular term is used not the plural. in German

7.3.2 Spellings

The most widely accepted spelling of words should be adopted for concept labels. If variant spellings exist and are commonly recognized, each term should be entered in the thesaurus with an appropriate attribute type. Optionally, references can be made to the preferred form, as in the examples below.

EXAMPLE 1

In English:		
<i>Roumania</i>	USE	Romania
<i>Rumania</i>	USE	Romania
In French:		
<i>Lithuanie</i>	EM	Lituanie

EXAMPLE 2

In English:		
non-fiction	USE	nonfiction
In French:		
co-voiturage	EM	covoiturage

Where possible, spelling should follow the practice of a well-established dictionary or glossary in the language of the thesaurus. If a choice between spellings is made for reasons of dialect, the chosen source should be stated in the introduction to the thesaurus (see [14.4](#)), and the choice should be adhered to consistently throughout.

However, proper names can cause legitimate inconsistencies in spelling practice, as they should be spelt according to the practice of their owners or recognized authorities such as the Virtual International Authority File (VIAF). See Example set 3, which assumes British spelling for terms other than proper names as appropriate.

EXAMPLE SET 3

colour	UF	<i>color</i>
defence	UF	<i>defense</i>
Department of Defense		(a US government department)
Ministry of Defence		(a UK government department)

Misspelt words are not acceptable in concept labels. However, some misspellings are so common that they can provide useful entry points, especially in electronic thesauri, and they may be used as alternate terms.

EXAMPLE 4

In English:		
abattoirs		
UF	<i>abatoirs</i>	
	<i>abbatoirs</i>	
	<i>abbattoirs</i>	
In French:		
Mitterrand François		
EP	<i>Mitterand François</i>	

Misspelt entry points should be provided only where:

1. the misspelling is common enough to be worthwhile; and
2. there is no danger that the misspelling relates to a term other than the preferred term indicated.

A term attribute type specifically for misspellings should be used and a tag such as MS (meaning MisSpelling).

When misspelt words are entered as alternate terms, a choice may be made to suppress the misspelt entries in printed or electronic displays, where they add nothing to a user's understanding of the concept but retain them in electronic versions used to interpret search terms entered by users.

7.3.3 Loan terms and translations of loan terms

Loan terms may be used as preferred terms if they are well-established. (See also [10.3.3.3](#).)

EXAMPLE SET 1 (of loan terms)

In English: bouquets gestalt therapy ombudsmen
In French: handball permafrost pizza

Occasionally, a loan term and a putative translation coexist. If the loan term is more widely accepted, it should be treated as the concept label, whereas the translation may be preferred if it has become well-established. Reciprocal references should be made between the two terms.

EXAMPLE SET 2 (of loan terms with translations)

In English: <i>coiffeurs</i> USE hairdressers hairdressers UF <i>coiffeurs</i>
abattoirs UF <i>slaughterhouses</i> <i>slaughterhouses</i> USE abattoirs
In French: baladodiffusion EP <i>podcast</i> podcast EM <i>baladodiffusion</i>

7.3.4 Transliteration

When transliterating terms from languages with different alphabets and writing systems, a recognized scheme should be followed. The scheme used should be specified in the introduction to the thesaurus. Some relevant schemes can be found within the standards publications listed in the Bibliography. See also [13.4](#), which addresses machine encoding issues associated with use of alternative scripts.

7.3.5 Neologisms, slang terms and jargon

Established terms for concepts should in general be used instead of slang or jargon terms, since these are often transient. Useful neologisms, however, are often coined in response to emerging technologies or developments in society or politics. They may be adopted as thesaurus terms as soon as they are judged useful for the retrieval, management, and use of information. In other cases, neologisms may be appropriate where established terms are now considered culturally inappropriate (see [Section 14.3.5](#)). However, the previous term should be included as a synonym so that earlier works using the term can be found in

the literature indexed. Slang or alternative terms may be appropriate when a newly emerging concept is expressed by a term originating within a particular subculture or social group. Where an alternative term is widely used, it may serve as a useful entry term, or it may constitute the concept label if there is a strong user preference and no widely accepted established term (see [Section 14.3.5](#)).

7.3.6 Common names and trade names

A product is frequently known by a recognized trade name. Where a suitable common name also exists, this should normally be adopted as the concept label, and the trade name should be admitted as an alternate term only if it is likely to serve as a user's access point. As registered trademarks enjoy legal protection, the symbol "®" should be added to the term to avoid legal problems.

EXAMPLE 1

soluble coffee
UF Nescafé®
Nescafé®
USE soluble coffee

Exceptions can occur, for example in applications where one generic product is sold under several competing trade names, and the users want to distinguish between the different brands. In this case the concept label for the generic product should be the common name, and the trade names may be used in narrower concepts. See also the discussion of specificity in 0.

Example, "Gramophone" was trademarked 1887-1910, but now "gramophone" is a generic name for the type of object.

7.3.7 Popular names and scientific names

If a popular and a scientific name refer to the same concept, preference should be given to the form more likely to be sought by users of the thesaurus. For example, "penguins" might be chosen as the concept label in a general thesaurus, but the scientific equivalent, "Sphenisciformes", may be preferred in a zoological thesaurus. Reciprocal references between the terms should be made in these cases with appropriate attribute types.

7.3.8 Abbreviations and acronyms

Abbreviations and acronyms are often ambiguous, because they can refer to more than one concept. Therefore, the full form of the name should normally function as the concept label or preferred term, with a reciprocal reference from the abbreviated form.

EXAMPLE SET 1 (of acronym/abbreviation disambiguation)

In English: United Nations Environment Programme UF <i>UNEP</i> <i>UNEP</i> USE United Nations Environment Programme direct current UF <i>DC (direct current)</i> <i>DC (direct current)</i> USE direct current national insurance UF <i>NI (national insurance)</i> <i>NI (national insurance)</i> USE national insurance	In French: nouveaux pays industriels EP <i>NPI (nouveaux pays industriels)</i> <i>NPI (nouveaux pays industriels)</i> EM nouveaux pays industriels
---	---

Exceptionally, abbreviations or acronyms may be used as concept labels when they are widely known, unambiguous and readily understood within the field covered by the thesaurus, especially if they have become so well-established that the full form of the name is rarely used or is generally ignored. Reciprocal references should still be made between the full term and its abbreviation.

EXAMPLE SET 2 (of acronyms/abbreviations which do not need disambiguation)

In English: UNICEF UF <i>United Nations International Children's Emergency Fund</i> <i>United Nations International Children's Emergency Fund</i> USE UNICEF HIV UF <i>human immunodeficiency virus</i> <i>human immunodeficiency virus</i> USE HIV	In French: SIDA EP <i>Syndrome d'immunodéficience acquise</i> <i>Syndrome d'immunodéficience acquise</i> EM SIDA
---	---

7.3.9 Proper names

7.3.9.1 General

Proper names are often useful in a thesaurus but are sometimes excluded if there are too many candidates, especially if their forms are controlled by another authority list. However, they should be included if the thesaurus is the only means of validating index terms. Inclusion as concepts has the additional benefit that

hierarchical or associative links (see [Clause 11](#)) can be established between concepts and relevant proper names, for example when the latter represent instances of the former.

EXAMPLES

mountains	
NT	Ben Nevis Mount Everest
Hockney, David (1937-)	
BT	artists
RT	contemporary art

7.3.9.2 Place names

Names of countries and geographical regions sometimes vary from language to language. Variant terms referring to the same place also occur within a single language community for reasons such as the following:

1. a "popular" and an "official" name are both in common use;

EXAMPLE 1

Republic of Korea South Korea

2. the vernacular form of the name differs from the usual form of the name in the language of the thesaurus;

EXAMPLE 2

Leghorn (A thesaurus in English may include this translation of the Italian city locally known as "Livorno") Livorno (The same thesaurus may also include the vernacular name)

3. in a country with more than one official language, two forms may commonly be used;

EXAMPLE 3

Gand (In Belgium, this French name for the city of Ghent coexists with the name in Dutch) Gent (In Belgium, this Dutch name for the city of Ghent coexists with the name in French)
--

The name that is most familiar to the users of the thesaurus should be designated as the concept label. Preference should be given to the official rather than the popular name. The short form of the official name should be preferred. Standard authorities, such as ISO 3166, should be consulted for the official forms. Reciprocal references can be made between the preferred and alternate forms.

EXAMPLE 4

Ireland
UF <i>Irish Republic</i>
<i>Eire</i>
<i>Eire</i>
USE Ireland
<i>Irish Republic</i>
USE Ireland

7.3.9.3 Names of institutions and persons

Variant names are common and can create problems if they are not controlled. When included in a thesaurus, the form of the names should be selected in accordance with a recognized code of cataloguing practice or manual of style. Care should be taken to ensure the name used as the concept label (preferred term) is correct, current and sufficiently complete to avoid confusion with other persons or institutions. Alternative forms of the name that are in common use should be established as alternate terms.

EXAMPLES

Armstrong, Louis (1901-1971)	UF	Satchmo (1901-1971)
Berners-Lee, Tim		
Otlet, Paul (1868-1944)		
Centre national de la danse (France)	UF	CND (Centre national de la danse)
Intergovernmental Committee for Physical Education and Sport		
Association for Professional Broadcasting Education (U.S.)	UF	Broadcast Education Association (U.S.)

7.4 Clarification and disambiguation of thesaurus terms

7.4.1 General

In ordinary discourse, a single term might have more than one meaning, and the most common meaning is not always the one needed to represent the concept wanted in the thesaurus. If the context provided by the hierarchy of broader and narrower concepts linked to the concept in question is insufficient to elucidate the intended scope, additional measures should be applied. While all thesaurus terms should be expressed as unambiguously as possible, it is especially important to formulate the concept label of a given concept in such a way that it conveys the intended scope to any user. For example, the multi-meaning term "depression" could be reformulated as "economic depression" or "meteorological depression", as appropriate. Alternatively, a qualifier (see [7.4.2](#) here) may be used. In cases where these measures are not appropriate or sufficient, or where additional information would help to clarify the meaning and make the usage more consistent, an explicit scope note should be used (see [6.2](#)).

7.4.2 Homographs and qualifiers

Homographs (sometimes referred to by the broader terms homonyms) are words with the same spelling but different meanings. Disambiguation is the process of making a clear distinction between two words which look the same but have different meanings depending on the context in which the word is used.

EXAMPLE 1

cranes (This term can refer either to birds or to lifting equipment.)

When homographs are needed as thesaurus terms, the meaning of each term should be clarified and the traditional way to do this is by adding to it a qualifier in parentheses. The qualifier should be as brief as possible, ideally consisting of one word. Often a term used for a broader concept, the qualifier should indicate

the context or subject area to which the concept belongs. It forms part of the term and does not serve as a scope note (see 6.2).

EXAMPLE 2

cranes (birds) cranes (lifting equipment)
--

A qualifier should be added to each homographic term, even when one of its senses is more common than the other in the domain of the thesaurus. For example, in English, "beams (structures)" is an appropriate preferred term in an engineering thesaurus that also includes "beams (radiation)". For example, in French, "eau (boisson)" is an appropriate preferred term in an environmental thesaurus that also includes "eau (environment)".

If only one of the homographic terms is used in a specialized thesaurus, and its meaning is obvious to the users of the thesaurus, the qualifier may be left out. However, consideration should be given to the possibility that the thesaurus scope may be expanded in the future, or that interoperability with other vocabularies covering different scopes might be needed.

Since the qualifiers make the terms slightly cumbersome to apply, and since some thesaurus applications (for example automated indexing) have difficulties in applying them, their use (especially in concept labels) should be avoided if another means of resolving ambiguity can be found. For this reason, a multi-word term should be preferred to a single-word term with a qualifier, as long as the compound form occurs in natural language.

EXAMPLE 3

For example, "industrial plants" should be preferred to "plants (facilities)". The latter expression, however, might be entered as an alternate term.

Parenthetical qualifiers should not be used to provide a term in inverted form.

EXAMPLE 4

For example, the inverted forms "cooking (fish)" and "pens (fountain)" are incorrect and the direct forms "fish cooking" and "fountain pens" should be used instead. "Fountain" is used in the latter example to indicate a type of pen rather than to disambiguate the word "pen". (See also 0 on the order of words in multi-word terms). Appropriate uses of qualifiers with the term "pens" in English are: "pens (enclosures)" and "pens (writing implements)". Appropriate uses of qualifiers with the term "audition" in French are: "audition (physiologie)" and "audition (communication)".
--

In the case of acronyms and abbreviations, which are often homographs, the qualifier should be the full form of the term (see also 7.3.8 Abbreviations and acronyms). Qualifiers for acronyms and abbreviations should only be used when their meaning is ambiguous to the users of the thesaurus, i.e. when they have another known meaning either in the domain covered by the thesaurus or in common language.

7.4.3 Definitions

A full definition is not usually required to clarify the way in which a term should be used. However, if a definition is required for another reason, a separate note field should be established for the definitions so that they do not become confused with any scope notes. The source of each definition should be recorded alongside the definition itself.

EXAMPLE

chiaroscuro DEF The style of pictorial art in which only the light and shade are represented (OED)
--

NOTE "OED" refers to the Oxford English Dictionary (www.oed.com). Any such acronyms in thesaurus definitions or scope notes should be explained in the introduction to the thesaurus (see 0).

7.4.4 History notes

When a thesaurus is updated, changes in terms can affect the retrievability of a concept. In such cases a history note may be used.

A history note may apply to a single preferred or alternate term, or to a concept. It should be used when a new preferred term is added to the thesaurus, or a change is made to an existing term that affects the concept's scope in different periods of application. While it is possible to include such information in the scope note, a separate history note is preferable. A history note may record the date of introduction, or it may give more complex advice on how to search for the concept in earlier or later times.

EXAMPLE 1

microwave ovens HN Concept introduced 1985
--

EXAMPLE 2

notebook computers HN Term introduced 1999; prior to that use "laptop computers" undocumented immigrants HN: Formerly "illegal aliens," deprecated in 2023
--

7.5 Grammatical form of terms

7.5.1 Nouns and noun phrases

A thesaurus term should preferably consist of a noun or a noun phrase. This includes gerunds, which are verbal nouns. In the English language (and in others, such as German, French and other romance languages) noun phrases occur in two forms:

1. adjectival phrases (including some where a noun or possessive functions as adjective)

EXAMPLE SET 1

Burkitt's lymphoma cold fusion environmental luminosity stone walls tropical diseases

2. prepositional phrases

EXAMPLE SET 2

accessories after the fact prisoners of war barrel vault with lunette_fat over lean technique [painting technique]
--

As prepositions can unnecessarily add to the length and clumsiness of a term, they should be avoided, if possible, especially in preferred terms. For example, use "carbohydrate metabolism" rather than "metabolism of carbohydrates". (However, the latter term could still be admitted as an alternate term.)