SLOVENSKI STANDARD
oSIST ISO/PRF 26162-3:2022

01-december-2022

Upravljanje terminoloških virov - Terminološke baze podatkov - 3. del: Vsebina

Management of terminology resources — Terminology databases — Part 3: Content

Gestion des ressources terminologiques — Bases de données terminologiques — Partie 3: Contenu

Ta slovenski standard je istoveten z: **ISO/PRF 26162-3:2022**

ICS:

01.020	Terminologija (načela in koordinacija)	Terminology (principles and coordination)
01.140.20	Informacijske vede	Information sciences
35.240.30	Uporabniške rešitve IT v informatiki, dokumentiranju in založništvu	IT applications in information, documentation and publishing

oSIST ISO/PRF 26162-3:2022 **en,fr,de**

INTERNATIONAL STANDARD

ISO 26162-3

First edition

Management of terminology resources — Terminology databases —

Part 3: Content

Gestion des ressources terminologiques — Bases de données terminologiques —

Partie 3: Contenu

PROOF/ÉPREUVE

Reference number
ISO 26162-3:2022(E)

COPYRIGHT PROTECTED DOCUMENT

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

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

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

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

This document was prepared by Technical Committee ISO/TC 37, *Language and terminology*, Subcommittee SC 3, *Management of terminology resources*.

A list of all parts in the ISO 26162 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

Managers, educators and terminology database maintenance authorities conduct both periodic and continuous evaluation of terminology databases containing concept entries for a number of purposes:

— quality-assurance-related validation of terminological data collections in business, government and non-governmental organizations;

— formative assessment and summative evaluation and feedback in training and educational environments.

ISO 26162-1 and ISO 26162-2 specify design principles and software considerations for modelling terminology databases (termbases). ISO 26162-1 establishes the general principles of termbase design as outlined in core ISO/TC 37 standards, such as ISO 704, which, among other topics, treats general principles for concept entry content and structure, term identification, basic principles for modelling concept systems and a range of other areas associated with terminology work. This document also encourages conformity to the terminological metamodel as outlined in ISO 16642. It describes the role that data categories play in modelling terminological data and sets down basic principles for ensuring and evaluating the quality of data stored in termbases, such as data granularity, elementarity and modelling variance. These criteria comprise fundamental benchmarks against which to measure the quality and reliability of terminological data. ISO 26162-2 relates the principles outlined in ISO 26162-1 to the implementation of database design with respect to software and user interface considerations, together with pragmatic workflow implementations in terminology management environments.

This document provides guidance for defining procedures for ensuring high-quality content in terminological data collections designed to meet documentation needs in a range of environments involving, for instance, translation, interpreting and technical communication. Conformity to this document can strengthen processes designed to support a quality management system, such as ISO 9001, and the related auditing procedures in a translation, interpreting or technical communication environment. An error typology is presented in the framework of an overall evaluation model, with generic (non-standardized) options for creating a concept entry evaluation model, depending on the needs of users and of the sponsoring organization.

Annexes A to C provide pragmatic advice on error evaluation practice. Annex A describes the creation of scoring models reflecting the error typology described in Clause 6, allowing for design variations depending on organization needs. For instance, a given scoring model can form the basis for a score card used for students and trainees, which is likely to be different from a score card used for a major enterprise or a national term bank.

Annex B presents a sample term entry. Annex C presents a sample evaluation model that can be adopted or adapted to meet the needs of terminologists, individuals working as freelancers or in companies, governmental organizations and NGOs. The values in this evaluation model can be used to create a scoring method, with the understanding that actual scoring practice is likely to vary according to specifications and user needs.

Management of terminology resources — Terminology databases —

Part 3: Content

1 Scope

This document specifies content-related aspects of terminology database maintenance. It gives guidance on the content of terminological data collections, with emphasis on data quality evaluation.

This document gives guidance for modellers of concept entries who need to ensure interoperability and high-quality content. It aims to ensure that terminological data collections themselves meet high standards for design conformity with standards such as ISO 12620-1 and ISO 16642, data accuracy and performance. It outlines principles for assuring data quality (see ISO 9001) and evaluating terminological data collections for purposes of continuous improvement. This approach contrasts that of ISO 23185:2009, which focuses on the usability of existing terminology resources.

This document does not apply to the management of text corpora or to term extraction tools.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 1087, *Terminology work and terminology science — Vocabulary*

ISO 26162-1, *Management of terminology resources — Terminology databases — Part 1: Design*

ISO 26162-2:2019, *Management of terminology resources — Terminology databases — Part 2: Software*

3 Terms and definitions

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

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
concept entry
entry
terminological entry
part of a *terminological data collection* (3.14) which contains the terminological data related to one concept

Note 1 to entry: A concept entry can contain information treating one or more languages.

Note 2 to entry: In this document, the term *entry* is used as a short form for *concept entry.*

[SOURCE: ISO 30042:2019, 3.5, modified — “entry” made second preferred term, notes to entry added.]

3.2
evaluation model
model for analysing data in a *concept entry* (3.1) according to terminology principles and specified data requirements consistent with the purpose of the *termbase* (3.11)

3.3
core structure
common structure and *data categories* (3.6) that are used in all TermBase eXchange (*TBX*) *dialects* (3.5)

Note 1 to entry: The core structure is compliant with ISO 16642 (TMF).

[SOURCE: ISO 30042:2019, 3.6]

3.4
error typology
systematic list of error types

3.5
TBX dialect
extensible markup language (XML) that validates according to the *core structure* (3.3) of TermBase eXchange (TBX) and allows exactly those *data categories* (3.6) at those levels specified by a precisely defined configuration of data categories

Note 1 to entry: See ISO 30042 for more detail.

[SOURCE: ISO 30042:2019, 3.12, modified — Simplified for this document, Note 1 to entry replaced.]

3.6
data category
class of data items that are closely related from a formal or semantic point of view

EXAMPLE /part of speech/, /subject field/, /definition/.

Note 1 to entry: A data category can be viewed as a generalization of the notion of a field in a database.

Note 2 to entry: In running text, such as in this document, data category names are enclosed in forward slashes, e.g. /part of speech/.

[SOURCE: ISO 30042:2019, 3.8]

3.7
data integrity
conformance of data values to a specified set of rules

[SOURCE: ISO/IEC TR 10032:2003, 2.23]

3.8
lexical unit
unit of language, belonging to the lexicon of a given language

[SOURCE: ISO 1951:2007, 3.8, modified — Reference to a dictionary removed.]

3.9
quality assurance
QA
set of planned and systematic activities necessary to provide confidence that a *concept entry* (3.1) satisfies acceptance criteria based on terminology principles and specified data requirements

3.10
specification
document that sets out detailed requirements to be satisfied by a *terminological data collection* (3.14)

Note 1 to entry: Specifications can include procedures for checking conformity to these requirements.

3.11
termbase
terminology database
database comprising a *terminological data collection* (3.14)

[SOURCE: ISO 30042:2019, 3.28, modified — "terminology database" is no longer an admitted term, but a second preferred term.]

3.12
termbase quality evaluator
person who is qualified as a terminologist or subject-matter specialist who conducts a quality evaluation of a *terminological data collection* (3.14)

3.13
termhood
degree to which a *lexical unit* (3.8) is recognized as a term

EXAMPLE "bulk carrier ship" has stronger termhood than "ship" alone. "Mouse" has termhood in computer applications, whereas it does not in general language.

Note 1 to entry: Termhood applies to both simple terms (consisting of a single word) and complex terms (consisting of more than one word or lexical unit), and to other designations, such as proper names and appellations, as well as formulas and symbols.

3.14
terminological data collection
resource consisting of *concept entries* (3.1) with associated metadata and documentary information

EXAMPLE A TBX document instance, ISO 1087.

[SOURCE: ISO 30042:2019, 3.29, modified — Second preferred term "TDC" removed.]

3.15
unithood

degree to which a given sequence of words has sufficient collocational strength to form a stable *lexical unit* (3.8)

EXAMPLE "art deco table" has stronger unithood than "modern table".

Note 1 to entry: Because unithood derives from the collocational relationship of words making up a given string, it only applies to multi-word terms.

4 Identifying terms

4.1 Requirements for term selection

Concept entries should meet the needs of their intended audience and purpose as well as organization-specific requirements, including requirements for terms. ISO 704 discusses principles for assigning and analysing designations.

While it is not possible to set universal requirements for term selection and for the content of concept entries, an important goal of a termbase in many commercial environments is prescriptive in nature: directing users away from terms that are problematic for one reason or another, and towards the use of preferred terms. Thus, term selection involves both identifying the organization's preferred terms and documenting the corresponding synonymous terms that are to be avoided. In this context implementers should document deprecated and obsoleted terms along with preferred terms, as well as non-central terms that nonetheless recur in critical enterprise content.