INTERNATIONAL STANDARD

ISO 26162-2

Management of terminology resources — Terminology databases —

Part 2:
Software

Gestion des resources terminologiques — Bases de données terminologiques —

Partie 2: Logiciels

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 26162-2:2019
https://standards.iteh.ai/catalog/standards/sist/3cbdbc37-b32d-4d71-a6ed-
6d2edc6b6f50/iso-26162-2-2019

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*.

This first edition of ISO 26162-2, together with ISO 26162-1, cancels and replaces ISO 26162:2012, which has been technically revised.

The main changes compared to the previous edition are as follows:

— the document has been split into parts. The first part is focusing on the design of terminology database design, the second part on the development of terminology management systems;

— all references to generic software design principles and specific use cases have been removed.

A list of all parts of 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

Terminologies are the totality of concepts in given subject fields represented by terms and other designations and described by using additional terminological data. In general, these data are organized in structured terminology databases and are usually manipulated in specific software applications called terminology management systems. Terminology databases usually vary with regard to their underlying data model and consist of different sets of data categories, while terminology management systems generally differ depending on their functionality and the platform they are designed for.

The ISO 26162 series gives guidance on designing terminology databases and on essential terminology management system features. The series can also be used to evaluate the conformance and suitability of terminology databases and terminology management systems.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 26162-2:2019
https://standards.iteh.ai/catalog/standards/sist/3cbdbc37-b32d-4d71-a6ed-
6d2edc6b6f50/iso-26162-2-2019

Management of terminology resources — Terminology databases —

Part 2:
Software

1 Scope

This document specifies essential features of terminology management systems, regardless of specific software engineering paradigms, user interface and user assistance design principles, and specific data models. These features enable maximum efficiency and quality in terminology work and, thus, support creating, processing, and using high quality terminology. The intended audiences of this document are software engineers/developers as well as terminologists, technical communicators, translators, interpreters, language planners, and subject field experts.

This document describes all features needed for recording, editing, maintaining, exchanging, and presenting terminological data. Term extraction features used to identify new terms are out of the scope of this document.

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 704, *Terminology work — Principles and methods*

ISO 1087, *Terminology work — Vocabulary*

ISO 8601 (all parts), *Date and time — Representations for information interchange*

ISO 12620, *Management of terminology resources — Data category specifications*

ISO 16642, *Computer applications in terminology — Terminological markup framework*

ISO 23185, *Assessment and benchmarking of terminological resources — General concepts, principles and requirements*

ISO 26162-1:2019, *Management of terminology resources — Terminology databases — Part 1: Design*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 1087, ISO 26162-1 and the following apply.

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

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

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

3.1
terminology management system
TMS
software tool specifically designed with a metadata structure for collecting, maintaining, and accessing terminological data

[SOURCE: ISO 1087:2019, 3.6.13]

3.2
concept entry structure
part of a terminology database that defines the structure for concept entries

3.3
input template
template that is created based on a *concept entry structure* (3.2) (or a subset of it) for the purpose of data input

Note 1 to entry: Depending on the degree of differentiation of user roles, more than one input template can be necessary.

3.4
layout template
template that defines how concept entries are displayed in the user interface

Note 1 to entry: Depending on the degree of differentiation of user roles, more than one layout template can be necessary.

3.5
terminology database definition
database information that defines the data categories and the *concept entry structure* (3.2) of a terminology database

Note 1 to entry: Depending on the degree of management requirements, more than one terminology database definition can be necessary.

4 Terminology management systems

4.1 General

The design of terminology management systems requires a deep understanding of terminology theory and terminology work. In this sense, and in order to achieve high quality results, the following shall be used:

— established terms and definitions as defined in ISO 1087;

— principles and methods as defined in ISO 704;

— data modeling criteria as defined in ISO 16642 and ISO 12620;

— usability metrics as defined in ISO 23185.

Terminology databases are used for collecting terminologies. Terminologies can achieve a high level of structural complexity and can be created and used by different user groups, each of them with their own specific needs. Consequently, terminology management systems shall be designed to respond to a wide range of potential use cases that can differ from those of conventional database management systems, particularly with regards to human-machine interaction and the interaction with other terminology-processing systems, such as translation memory systems, localization tools or authoring tools.

4.2   Terminological metamodel

Every terminology database shall comply with the terminological metamodel defined in ISO 16642 (for more information see also ISO 26162-1). Thus, terminology management systems shall allow for the instantiation of every level of the terminological metamodel, from high-level containers (global and complementary information) down to the term component section.

4.3   Creating terminology databases

4.3.1   Terminology database definition

When creating terminology databases, the terminology management system shall support the following essential features to allow users to perform the following actions:

— assign a name to a terminology database;

— describe a terminology database (content, rights, copyright, etc.);

— provide a core concept entry structure based on the terminological metamodel (see 4.2) including the option to add/remove the term component section;

— use predefined concept entry structures;

— set up customizable concept entry structures;

— aggregate data categories at the core concept entry structure and further nest these data categories;

— explicitly aggregate data categories[1], such as /definition/, /source/ or /note/ at the language level;

— replicate part of the concept entry structure when setting up a terminology database (for example, copy a set of data categories from the French language section for creating an identical data category set for the Spanish language section);

— repeat data categories (see ISO 26162-1:2019, 4.4.3);

— define the use (e.g. mandatory, see ISO 26162-1:2019, 4.3) and the cardinality of data categories (see 4.3.3);

— preview terminology database definitions;

— extend, condense and/or otherwise change terminology database definitions.

4.3.2   Data categories

Terminological data in specific data fields are instances of terminological data categories that have been previously defined during the terminology database setup and then have been associated with the corresponding levels of the terminological metamodel (for data categories see ISO 26162-1:2019, 4.3). The terminology management system shall also allow parent-child relationships between data categories, such as the relationship between /definition/ and /source/ illustrated in the following example:

EXAMPLE

Concept entry identifier: 3 (for concept *circuit-breaker*)

  Concept position: 1.2

---

1)   In this document, data category names used in running text are indicated between forward slashes (for example, /definition/).

Definition: <u>mechanical switching device</u>, capable of making, carrying and breaking currents under normal circuit conditions and also making, carrying for a specified duration and breaking currents under specified abnormal circuit conditions such as those of short circuit[2)]

Source (of definition): <u>Electropedia</u>

During terminology database setup, terminology management systems shall provide a predefined set of data categories and picklist values from established repositories and standards in order to support a high grade of interoperability between terminology databases (and other applications and resources). For example, common terminological data categories are specified in DatCatInfo (see Reference [10]), language codes are defined in ISO 639 (all parts), country codes are defined in ISO 3166 (all parts), and subject field codes can be taken from universal classification systems, such as the Universal Decimal Classification. However, terminology management systems shall also allow for the creation of customizable data categories and picklist values if appropriate data categories or picklist values are not available. If predefined data categories and picklist values are implemented for use in multilingual environments, these data categories and picklist values shall be made available in the desired user interface language.

4.3.3 Data types and cardinalities

In terminology, free text strings (such as terms, definitions or contexts), picklists, binaries (such as figures, audio or video), cross-references (such as between concept entries or links to external resources), and Boolean values are the most common data types. However, in order to cover any potential use case, terminology management systems should support other data types like those defined for XML Schema, for example, or other schema or programming languages.

Furthermore, terminology management systems shall allow users to configure cardinalities, i.e. to set a specific number of occurrences for any data category, for instance, to allow for only one definition at the concept entry level.

4.3.4 Roles and rights

Distributed terminology work usually requires the assignment of different roles and rights (see also ISO 26162-1:2019, 4.3.2.3 and 4.5). Specific terminology database objects (such as terminology database definitions, input templates, export/import profiles, filters) or instances (such as concept entries, certain language and/or certain term sections) require specific roles and rights. Thus, appropriate roles and rights can be granted depending on:

— authorization types, such as supervisor, creator, proposer, end user;

— work profiles, such as author, translator, service and marketing staff.

4.4 Data input and editing

In the context of data input and editing, terminology management systems shall support fundamental features to allow users to perfom the following actions:

— use predefined input templates;

— set up customizable input templates;

— copy parts of input templates for reuse;

— preview input templates during setup;

— copy concept entries (or parts of concept entries) to be used as a basis for the creation of new concept entries (or for the modification of concept entries, for example, by inserting a new language section copied from an adjacent concept entry);

---

2) Definition taken from Electropedia, http://www.electropedia.org

— automatically generate administrative data, such as authorship for creating and updating a concept entry or an instance of it, as well as concept entry identifiers and time/date stamps, according to ISO 8601 (all parts);

— set certain values for certain data fields when creating an input template. For example, /part of speech/ could have the default value "noun" if it is predetermined that most terms in the terminology database will be nouns. This can increase productivity by not requiring users to select the value for /part of speech/ for every term. However, users shall have the option to change these values at any time. To prevent unintentional contamination of data, it is recommended to prompt users to review their default values when starting a new session;

— automatically propagate changes to data elements throughout the whole terminology database or instances of it. For example, data fields can contain misspelled words or inconsistent spelling that need to be tackled by search and replace features. The need for search and replace can also apply to picklist values once they have been changed in the terminology database definition;

— add links to other concept entries, such as a link to a concept entry for an entailed term that is used in a definition or other text field, even if the term in the text does not have the same surface form as the term that designates a concept in another concept entry. For example, it shall be possible to link the plural form "circuit-breakers" in a definition to the concept entry that contains the term "circuit-breaker" as a designation of the corresponding concept. Terminology management systems shall offer the option to establish links between concept entries using terms or concept entry identifiers;

— add links to external resources (multimedia files, web pages, etc.);

— define as many languages as needed (multilingual environments);

— save a search term as candidate for a new concept entry when a search has returned no results;

— alert the user when trying to introduce a term that already exists. However, it shall be possible to enter duplicate terms in order to accommodate the existence of homographs;

— save partially completed concept entries in draft mode, and then allow additional information to be added as it becomes available;

— perform bulk editing and deletion of concept entries.

4.5   Searching and filtering

Most typical users are interested in finding information about a term or other kind of terminological information within concept entries or throughout the whole terminology database. In particular, terminology management systems shall allow users to:

— look up the definition of a concept;

— search for a target language equivalent;

— confirm a term's spelling or check the contexts in which it is used;

— look up whether a term has abbreviations or synonyms or look up the full form of an abbreviation;

— display all terms representing a concept including their status;

— find information about a concept and its relationships to other concepts;

— use all languages as search languages;

— apply predefined search and filter criteria;

— carry out complex searching or filtering for data management purposes, for example, to find concept entries that have been updated on a specific date or by a specific person, to find duplicate concept entries or to find concept entries where specific fields are empty.