TECHNICAL SPECIFICATION

ISO/TS 24634

First edition
2021-11

Management of terminology resources — TBX-compliant representation of concept relations and subject fields

Gestion des ressources terminologiques — Représentation des relations conceptuelles et des domaines conforme à TBX

Reference number
ISO/TS 24634:2021(E)

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*.

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

This document describes best practices for specifying subject fields and concept relations in terminology databases (termbases). It also demonstrates how to represent subject fields and concept relations in terminological document instances in a way that is compliant with ISO 30042.

Concept relations for specific TBX dialects are specified in the form of dedicated TBX modules. Subject fields can be declared in the TBX backmatter or implemented through an XML namespace. This document is intended to maximize interoperability of these types of information.

Throughout this document, reference is made to data categories (DCs). To maximize interoperability, it is essential that termbases use the same DCs, as described in this document, for the same purposes. DatCatInfo[2] is a publicly available electronic repository of data category specifications. Data categories used in the examples in this document are taken from DatCatInfo.

This document complements ISO 30042.

Management of terminology resources — TBX-compliant representation of concept relations and subject fields

1 Scope

This document provides requirements and recommendations for representing subject fields and concept relations in TBX-compliant terminological document instances. Examples in this document utilize the DCA style of TBX markup.

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 12620-1[1)], *Management of terminology resources — Data categories — Part 1: Specifications*

ISO 12620-2[2)], *Management of terminology resources — Data categories — Part 2: Repositories*

ISO 30042, *Management of terminology resources — TermBase eXchange (TBX)*

W3C, *SKOS Simple Knowledge Organization System Reference*, W3C Recommendation 18 August 2009. Available at: https://www.w3.org/TR/skos-reference/

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 1087, ISO 12620-1, ISO 12620-2, ISO 30042 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 relation
relation between concepts

[SOURCE: ISO 1087:2019, 3.2.11]

3.2
hierarchical relation
hierarchical concept relation
generic relation (3.3) or *partitive relation* (3.4)

[SOURCE: ISO 1087:2019, 3.2.12]

1) Under preparation. Stage at the time of publication: ISO/DIS 12620-1:2021.

2) Under preparation. Stage at the time of publication: ISO/DIS 12620-2:2021.

3.3
generic relation
generic concept relation
genus-species relation
concept relation (3.1) between a generic concept and a specific concept where the intension of the specific concept includes the intension of the generic concept plus at least one additional delimiting characteristic

[SOURCE: ISO 1087:2019, 3.2.13, modified — Example and Notes to entry omitted.]

3.4
partitive relation
partitive concept relation
part-whole relation
part-of relation
concept relation (3.1) between a comprehensive concept and a partitive concept

[SOURCE: ISO 1087:2019, 3.2.14, modified — Example omitted.]

3.5
associative relation
associative concept relation
pragmatic relation
non-hierarchical *concept relation* (3.1)

[SOURCE: ISO 1087:2019, 3.2.23, modified — Example omitted.]

3.6
concept entry
terminological entry
part of a terminological data collection which contains the terminological data related to one concept

[SOURCE: ISO 30042:2019, 3.5, modified — Term "entry" omitted.]

3.7
data category
DC
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 categories are enclosed in forward slashes (e.g. /part of speech/).

[SOURCE: ISO 30042:2019, 3.8]

3.8
picklist
list of permissible values of a closed *data category* (3.7)

3.9
subject field
domain
field of special knowledge

[SOURCE: ISO 10241-1:2011, 3.3.1, modified — "subject field" is the preferred term, and Notes to entry omitted.]

3.10
subject-field classification
organization of the *subject fields* (3.9) and subfields dealt with in a *terminological data collection* (3.12) into a logical structure

3.11
termbase
terminology database
database comprising a *terminological data collection* (3.12)

[SOURCE: ISO 30042:2019, 3.28]

3.12
terminological data collection
resource consisting of *concept entries* (3.6) with associated metadata and documentary information

EXAMPLE A TBX document instance.

[SOURCE: ISO 30042:2019, 3.29, modified — Admitted term “TDC” omitted, and example “ISO 1087” omitted.]

4 Subject fields

4.1 General

Terminology databases (termbases), particularly large ones, frequently need to be organized according to subject fields. In this case, there may be a /subject field/ data category at the concept level (this is implemented as a `<descrip>` element).

A picklist shall be used as the content model for subject fields. If the organization in question has a taxonomy that reflects its field of activity, the subject-field classification should reflect that taxonomy if appropriate.

Subject-field classifications can be declared in the backmatter of a TBX document instance, or through an XML namespace. In this document, the backmatter approach is described. The XML namespace approach requires a DCT style of TBX markup and can be modelled in parallel to the backmatter method.

A given TBX document instance can use more than one subject-field classification.

4.2 Specifying the name of the subject-field classification

The name of the subject-field classification used in a TBX document instance shall be declared in the TBX header.

EXAMPLE 1

```
<tbxHeader>
   <fileDesc>
      <sourceDesc>
         <p type="source">Termbase from ABC company</p>
         <p type="subjectFieldClass" id="ABC-Class">ABC Subject-Field Classification</p>
      </sourceDesc>
   </fileDesc>
</tbxHeader>
```

Additional information about the subject field shall be provided in the backmatter, as described in subsequent sections of this document. For this purpose, the `id` attribute acts as a unique pointer to the relevant section in the backmatter.

When more than one subject-field classification is used in a TBX document instance, the additional name declaration shall be provided.

EXAMPLE 2

```
<tbxHeader>
   <fileDesc>
      <sourceDesc>
         <p type="source">Termbase from ABC company</p>
         <p type="subjectFieldClass" id="ABC-Class">ABC Subject-Field Classification</p>
         <p type="subjectFieldClass" id="DEF-Class">DEF Subject-Field Classification</p>
      </sourceDesc>
   </fileDesc>
</tbxHeader>
```

If the additional subject-field classification is a result of the merging of two termbases, the name of the termbase may also be provided.

EXAMPLE 3

```
<tbxHeader>
   <fileDesc>
      <sourceDesc>
         <p type="source">Termbase from ABC company</p>
         <p type="subjectFieldClass" id="ABC-Class">ABC Subject-Field Classification</p>
         <p type="source">Termbase from DEF company</p>
         <p type="subjectFieldClass" id="DEF-Class">DEF Subject-Field Classification</p>
      </sourceDesc>
   </fileDesc>
</tbxHeader>
```


4.3 Defining the scope of subject-field values

Terminologists and other users of termbases often find it difficult to determine which subject field a concept should be assigned to. This is largely because historically the scope and meaning of subject fields themselves have not been defined for the users. Therefore, it is recommended that a clear description of the scope of each subject field be available to users of the termbase. In this document, a method is described for recording this information in the backmatter of a TBX document instance.

Some termbases use a publicly available subject-field classification, such as EuroVoc[3] or Lenoch[4]. These sources provide descriptions of the scope and meaning of their subject fields, and therefore, it is recommended to avoid duplicating this information in a termbase.

Termbases that adopt a unique subject-field classification should include information about the scope of the subject-field values in the backmatter of the TBX document instance.

In all cases, the subject-field description should be available or known to termbase users when they are assigning a subject-field value to a terminological entry. Figure 1 shows an example of a subject-field description from an existing termbase.

Subject field

- Start date: 2010-01-08
- End date:
- Number of occurrences: 2845
- Allow single language module: No
- Created by: user_131216
- Creation date: 2010-01-08 11:25:36
- Last modified by: user_8217
- Last modified date: 2012-09-14 11:56:20
- Code: BCO
- Notes:
  - 2012-01-16 : Added synonyms and notes (user_131216).
  - 2012-02-14 : Modified synonyms and added definitions (user_131216).
  - 2012-09-14 : Modified SF descriptor in the definitions (user_8217).

Languages

- English
  - Descriptor: Aquaculture
  - Synonyms: aquiculture; aquafarming; fish breeding; fish culture; fish farming;fish husbandry; fish rearing; fish spawning; shellfish farming; shellfishery
  - Definition: Terms related to both seawater farming (mariculture or sea farming) and freshwater farming, i.e. the cultivation of marine or fresh water aquatic plants or animals. Includes breeding or raising fish (fishfarming or pisciculture) and various shellfish such as shrimp, crayfish, oysters, mussels, scallops and aballone. Can also include pearl farming (pearl culture), seaweed farming (algal culture or algoculture) and coral farming (coral culture or coral gardening). Note that names of species are classified in subject fields under Botany (SF) or Zoology (SG).
- French
  - Descriptor: Aquaculture
  - Synonyms: aquiculture; culture marine; piscifacture; thalassoculture
  - Definition: Comprend la terminologie de l'élevage et de la multiplication des animaux et des plantes aquatiques, à des fins commerciales. Ce domaine regroupe la terminologie liée non seulement à l'aquaculture marine (mariculture), mais à l'élevage en eau douce (potamoculture). Comprend l'élevage de poissons (pisciculture ou élevage piscicole) et de crustacés tels que les crevettes (pénéiculture ou crevetticulture) ou les écrevisses (astaciculture), et l'élevage des coquillages (conchylicuture), notamment la culture des moules (mytiliculture), des huîtres (ostréiculture), des pétoncles ou des coquilles Saint-Jacques (pectiniculture) et des ormeaux (halioticulture). Peut également comprendre la culture d'algues (algoculture ou phycoculture), et la culture des perles (perliculture) ou du corail (culture de coraux ou coralliculture). Nota : Les noms d'espèces sont classés dans des domaines sous Botanique (SF) ou Zoologie (SG).
- Spanish
  - Descriptor: Acuicultura
  - Synonyms:
  - Definition:
- Portuguese
  - Descriptor: Aquicultura
  - Synonyms:
  - Definition:

Figure 1 — Sample description of a subject field

4.4 Hierarchy of subject fields

Frequently it is not sufficient to have a simple list of subject-field values without any parent/child relationships (referred to as a “flat” list). For subsetting and search purposes, large termbases benefit from a multi-level hierarchy of subject fields. Figure 2 shows a multi-level subject-field classification, with *Energy* having two subordinate levels, and *Environment and Natural Resources* having one subordinate level. Figure 3 shows another example from the field of mining.