

ISO/DTS 24634

ISO/TC 37/SC 3

~~Date: 2024-06-21~~

~~ISO/DTS 24634:2024(en)~~

Secretariat: DIN

Date: 2024-07-24

Management of terminology resources — TermBase eXchange (TBX)-compliant representation of concept relations and subject fields

Gestion des ressources terminologiques — Représentation des relations conceptuelles et des domaines conforme à TermBase eXchange (TBX)

iTeh Standards
(<https://standards.itih.ai>)
Document Preview

ISO/DTS 24634

<https://standards.itih.ai/catalog/standards/iso/d6ab1ec0-b2d8-4494-8cac-663be956257a/iso-dts-24634>

© ISO 2024

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
E-mail: copyright@iso.org
Website: www.iso.org

Published in Switzerland

iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

[ISO/DTS 24634](#)

<https://standards.iteh.ai/catalog/standards/iso/d6ab1ec0-b2d8-4494-8cac-663be956257a/iso-dts-24634>

Contents

Foreword.....	iv
Introduction.....	v
1 Scope.....	1
2 Normative references.....	1
3 Terms and definitions.....	1
4 Subject fields.....	3
4.1 General.....	3
4.2 Specifying the name of the subject-field classification.....	3
4.3 Defining the scope of subject-field values.....	4
4.4 Hierarchy of subject fields.....	5
4.5 Representing subject-field values in concept entries.....	6
4.5.1 General.....	6
4.5.2 Referencing the backmatter.....	7
4.5.3 Expressing the position of the subject-field value in each entry.....	7
4.6 Using a publicly available subject-field classification.....	7
5 Representing subject-field classifications in the backmatter.....	8
6 Concept relations.....	9
6.1 Typology of concept relations.....	9
6.2 Implementation of concept relations.....	10
6.2.1 General.....	10
6.2.2 TBX markup for concept relations within the concept entry.....	10
6.2.3 TBX markup for concept relations outside the concept entry.....	10
Annex A (informative) Typology of associative relations.....	12
Annex B (informative) Markup samples for concept relations in the <body> element.....	14
Annex C (informative) Markup samples of concept relations in the backmatter.....	17
Bibliography.....	19

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~~document 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).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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 second edition cancels and replaces the first edition (ISO/TS 24634:2021), which has been technically revised.

The main changes are as follows:

- ~~Addition~~addition of TBX markup for concept relations outside the concept entry in ~~6.2.3~~(6.2.3);
- ~~Typology~~typology of associative ~~concept~~relations in Table A.1 aligned with ISO 704:2022-(Annex A);
- ~~Markup~~addition of markup samples of concept relations in the backmatter in ~~Annex C~~(Annex C).

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 extensible markup language (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~~DC specifications. ~~Data categories~~DCs used in the examples in this document are taken from DatCatInfo^[2].

This document complements ISO 30042.

iTeh Standards (<https://standards.iteh.ai>) Document Preview

ISO/DTS 24634

<https://standards.iteh.ai/catalog/standards/iso/d6ab1ec0-b2d8-4494-8cac-663be956257a/iso-dts-24634>

Management of terminology resources — TermBase eXchange (TBX)--compliant representation of concept relations and subject fields

1 Scope

This document ~~providesspecifies~~ requirements and recommendations for representing subject fields and concept relations in TBX-compliant terminological document instances. Examples in this document utilize the data category as attribute (DCA) style (~~Data Category as Attribute~~) 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, *Management of terminology resources — Data categories — Part 1: Specifications*

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

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

W3C, *RDF Resource Description Framework*, W3C Recommendation 25 February 2014, Available at: <https://www.w3.org/RDF/>

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]

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~~notes to entry ~~omitted~~deleted.]

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~~deleted.]

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~~deleted.]

3.6

concept entry

terminological entry

part of a *terminological data collection* (3.12) which contains the terminological data related to one concept

[SOURCE: ISO 30042:2019, 3.5, modified — Admitted term “entry” ~~omitted~~deleted.]

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, modified — ~~The admitted—Abbreviated~~ term “DC” ~~has been~~added.]

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 replaced “domain” as~~ the preferred term, ~~and~~. Notes to entry ~~omitted~~~~deleted~~.]

3.10

subject-field classification

logical structure of the *subject fields* (3.9) and subfields dealt with in a *terminological data collection* (3.12)

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~~~~Abbreviated~~ term “TDC” ~~omitted~~, and example “ISO 1087” ~~omitted~~~~deleted~~.]

4 Subject fields

4.1 General

~~Terminology databases (termbases),~~ Termbases, particularly large ones, should 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. 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 data category as tag (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 [Clause 5](#) ~~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 can 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 should 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^[43] or Lench^[54]. 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 concept entry. [Figure 1](#) shows an example of a subject-field description from an existing termbase.