

---

---

**Systems to manage terminology,  
knowledge and content — TermBase  
eXchange (TBX)**

*Systèmes de gestion de la terminologie, de la connaissance et du  
contenu — TermBase eXchange (TBX)*

STANDARDSISO.COM : Click to view the full PDF of ISO 30042:2008



**PDF disclaimer**

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

STANDARDSISO.COM : Click to view the full PDF of ISO 30042:2008



**COPYRIGHT PROTECTED DOCUMENT**

© ISO 2008

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Published in Switzerland

## Contents

1 Scope.....	1
2 Normative references.....	1
3 Terms and definitions.....	2
4 Relationship to other standards.....	5
5 Applications of TBX.....	5
6 Fundamental principles.....	5
6.1 General.....	5
6.2 Principles relating to grouping and representing data-categories.....	6
7 Requirements for TBX files.....	7
7.1 Compliance requirements.....	7
7.2 Examples of non-compliance.....	8
7.3 Implementation levels.....	8
8 The core-structure module.....	8
8.1 Introduction.....	8
8.2 Hierarchy.....	9
8.3 Components of a terminological entry.....	10
8.4 Elements that can appear at multiple levels of the entry.....	10
8.5 Elements that occur only at the term level or lower.....	11
8.6 Handling of text.....	12
8.7 Meta data elements.....	14
8.8 Attributes.....	15
8.9 Character set issues.....	16
8.10 Language.....	16
9 The default data-category constraints.....	16
9.1 Introduction.....	16
9.2 Data-categories built into the core structure DTD of TBX.....	17
9.3 Data-categories specialized from meta data-categories through the default XCS file.....	17
10 Examples.....	21
10.1 Example of a typical TBX file.....	21
10.2 Examples of encoding TBX elements.....	22
10.3 Examples of TBX entries.....	23
11 Referencing objects.....	25
11.1 General information about referencing.....	25
11.2 Referencing a file that is embedded in the back matter of a TBX file.....	26
11.3 Referencing a file from the back matter.....	26
11.4 Referencing a file directly in the entry.....	27
11.5 Referencing an external source.....	27
11.6 Referencing and documenting a bibliographic source.....	27
11.7 Referencing and documenting information about a responsible person or organization.....	28
11.8 Referencing an external concept system, classification system, or thesaurus.....	29
11.9 Referencing a TBX entry from within a corpus.....	29
12 Creating customized TBX TMLs.....	29
12.1 General information about TMLs.....	29
12.2 Example of an XCS file for a user-defined TBX TML.....	30
12.3 Creating customized picklist display names.....	31
Annex A (Normative) DTD for the core structure module.....	33
Annex B (Normative) DTD for the data-category constraints (XCS file).....	38
Annex C (Normative) Default XCS file.....	40
C.1 Introduction.....	40
C.2 XCS file for the default data-categories and constraints.....	40
Annex D (Normative) Descriptions of the core structure elements and attributes and the default data-categories.....	48
Annex D.1 General information about the descriptions.....	48
D.2 Macros.....	48

D.3 Attribute classes.....	49
D.4 Elements.....	50
D.5 Default data-categories.....	61
Annex E (Normative) Descriptions of elements and attributes for the XCS file.....	73
E.1 Introduction.....	73
E.2 Attribute classes.....	73
E.3 Elements.....	73
Annex F (Informative) Integrated schema and other TBX resources.....	81
Annex G (Informative) TBX-Basic.....	82
Annex H (Informative) Summary of changes.....	83
Annex I (Informative) Indexes.....	88
I.1 Core-module DTD.....	88
I.2 XCS DTD.....	89
I.3 Terminological data-categories.....	89
Bibliography.....	91

STANDARDSISO.COM : Click to view the full PDF of ISO 30042:2008

## Foreword

The International Organization for Standardization (ISO) 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 ISO 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.

ISO 30042 was prepared by LISA OSCAR and was adopted, under a special "fast-track procedure", by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 3, *Systems to manage terminology, knowledge and content*, in parallel with its approval by the ISO member bodies.

The Localization Industry Standards Association (LISA - [www.lisa.org](http://www.lisa.org)) is the standards organization for the globalization industry. Within LISA, the OSCAR (Open Standards for Container/content Allowing Reuse) Special Interest Group develops XML-based standards for automated language-processing in the areas of globalization, internationalization, localization, and translation, including standards for translation memory, terminology, text memory, word/character counts, and other related areas. The main task of the OSCAR Special Interest Group is to develop standards to facilitate and automate the globalization of products and services in a way that supports local language and culture conventions. Publication as an OSCAR standard requires approval by the OSCAR steering committee. An earlier version of TBX was developed and published by LISA in 2002.

TBX and the TBX logo are registered trademarks of LISA, and the TBX logo is subject to terms of use as defined by LISA. LISA maintains copyright on the TBX specification that is available on the LISA Web site, and ISO maintains copyright on the TBX specification that it distributes as ISO 30042. The technical content of these two documents is identical, and is subject to joint maintenance by a team of ISO TC 37 and LISA OSCAR members.

## Introduction

This International Standard defines an XML-based framework for representing structured terminological data referred to as TermBase eXchange (TBX). Within this framework, a variety of terminological markup languages (TMLs) can be defined. A TML defined by TBX can facilitate the interchange of terminological data between users, which include people such as translators and writers, and applications and systems, such as Computer Assisted Translation tools and controlled authoring software. Therefore, it can be used for both human-oriented and machine-oriented terminological data. In this manner, it can enable the flow of terminological information throughout the information production cycle, both inside an organization and with outside service providers.

The intended audience for this document consists of two groups: (1) programmers and analysts who wish to develop software applications that process TBX-compliant data files; (2) terminologists and other language specialists who wish to analyse a terminological data collection for representation in TBX or to understand a TBX file.

This version of TBX is an update of a version that was published by the Localization Industry Standards Association (LISA) in 2002. Among other enhancements, the current version provides reference to an integrated schema that includes the core-structure module and the data-category constraints in combined declarations using the Relax NG and Schematron languages. It also provides reference to a TBX-compliant TML called TBX-Basic.

Users of this International Standard should first study the body (clauses 1-12). The suggested use of annexes A-I is described below.

### (1) The core-structure module of TBX

All TMLs within the TBX framework have the same core structure. The core-structure module is described in Clause 8. A DTD for the core-structure module is found in Annex A. The elements, attributes, and data types are described in Annex D, and listed alphabetically in Annex I.

### (2) The XCS module

TMLs may differ with respect to which data-categories are allowed, and at what levels of a terminological entry these data-categories can occur. These constraints on the core structure, which define a particular TML, are formally represented in an XCS file. A DTD for the XCS module is found in Annex B. The elements and attributes are described in Annex E, and listed alphabetically in Annex I.

### (3) The default XCS of TBX

The TBX-default TML is constrained by the default XCS file. The TBX default XCS is described in Clause 9. The default XCS file is provided in Annex C. The data-categories are described in Annex D, and listed alphabetically in Annex I.

### (4) Compliance checking of TBX document instances

Once a TBX TML has been defined by an XCS, a TBX document instance can be checked for compliance with that TML. The requirements for compliance are found in Clause 7. One can use a variety of methods and schema definition languages to check compliance. In particular, the Relax NG schema referred to in Annex F can be used to check whether a TBX document instance is compliant with the TBX-default TML. Annex F also indicates where a TBX user can find additional resources for compliance checking. Another TBX TML, called TBX-Basic, is referred to in Annex G.

### (5) Changes that have been made to TBX since its submission to ISO in February 2007 are summarized in Annex H.

Summary of annexes:

- A: DTD for core-structure module
- B: DTD for XCS module
- C: Default XCS that defines the TBX-default TML
- D: Descriptions of core structure elements and attributes
- D.5: Descriptions of default data-categories
- E: Descriptions of XCS elements and attributes
- F: Relax NG schema and other resources for compliance checking
- G: Reference to TBX-Basic
- H: Summary of changes to TBX
- I: Indexes (alphabetical lists of elements and data-categories)

# Systems to manage terminology, knowledge, and content - TermBase eXchange (TBX)

## 1 Scope

The TBX framework defined by this International Standard is designed to support various types of processes involving terminological data, including analysis, descriptive representation, dissemination, and interchange (exchange), in various computer environments. The primary purpose of TBX is for interchange of terminological data. It is limited in its ability to represent presentational markup. Intended application areas include translation and authoring.

TBX is modular in order to support the varying types of terminological data, or *data-categories*, that are included in different terminological databases (termbases). TBX includes two modules: a core structure, and a formalism for identifying a set of data-categories and their constraints, both expressed in XML. The term *TBX*, when used alone, refers to the framework consisting of these two interacting modules.

To maximize interoperability of the actual terminological data, TBX also provides a default set of data-categories that are commonly used in terminological databases. However, subsets or supersets of the default set of data-categories can be used within the TBX framework to support specific user requirements.

TBX, when used with its default set of data-categories, qualifies as a *terminological markup language (TML)* as defined in ISO 16642, which will be referred to as the *TBX-default TML* in this International Standard. Likewise, other markup languages that comply with TBX and use a subset of the default set of data-categories are also TMLs, but may go by other names, such as the one referred to in [Annex G \(Informative\) TBX-Basic](#).

## 2 Normative references

The following referenced documents are indispensable for the application 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 639-1:2002, *Codes for the representation of names of languages – Part 1: Alpha-2 code*

ISO 639-2:1998, *Codes for the representation of names of languages – Part 2: Alpha-3 code*

ISO 639-3:2007, *Codes for the representation of names of languages – Part 3: Alpha-3 code for comprehensive coverage of languages*

ISO/IEC 646:1991, *Information technology – ISO 7-bit coded character set for information interchange*

ISO 3166-1:2006, *Codes for the representation of names of countries and their subdivisions – Part 1: Country codes*

ISO 8601:2004, *Data elements and interchange formats – Information interchange – Representation of dates and times*

ISO/IEC 10646, *Information technology — Universal Multiple-Octet Coded Character Set (UCS)*

ISO 12200:1999, *Computer applications in terminology – Machine-readable terminology interchange format (MARTIF) – Negotiated interchange*

ISO 12620, *Computer applications in terminology – Data categories*

ISO 16642:2003, *Computer applications in terminology – Terminological markup framework*

## 3 Terms and definitions

For the purposes of this International Standard, the following terms and definitions apply:

**3.1 analysis**  
identification of the elements and structure of a terminological data collection in order to make explicit the data fields, their types, and their relationships

**3.2 blindness**  
property of a data format indicating the degree to which the data are sufficiently defined that it is unnecessary for the importer to establish contact with the originator of the data in order to interpret them

NOTE The term blindness has its origin in the engineering phrase "blind transmission," which refers to a transmission of data where it is not necessary to "see" who is the sender of the data in order to interpret it. In terminology, the concept of blindness is often used in the context of **blind interchange** (3.3).

**3.3 blind interchange**  
ability to receive a terminology file and integrate it into a target system, such as a Computer-Assisted Translation (CAT) tool, without having to contact the originator of the file in order to understand its contents

NOTE Interchange that is perfectly blind is interchange that is lossless while requiring no communication between the sender and the receiver of the data. Due to differences between terminological data collections and markup formats, perfectly blind interchange is rare. Typically, some of the data in a data collection is blind (can be exchanged without loss and without communication between parties) and some of the data requires communication between the parties in order to be exchanged.

**3.4 complementary information**  
CI  
information supplementary to that described in terminological entries and shared across the terminological data collection

[ISO 16642:2003]

NOTE In a TBX document instance, complementary information is contained in the back matter.

**3.5 core-structure module**  
XML specification of the elements and attributes that are permitted in a TBX file

NOTE The core-structure module is defined in a DTD which is used in tandem with an XCS file that applies additional data-category constraints. It can also be used to generate an integrated schema, such as a Relax NG schema [ISO 19757-2], that defines both the core-structure module and the data-category constraints in one file. See also **data-category constraint** (3.7).

**3.6 data-category**  
result of the specification of a given data field

[ISO 1087-2:2000]

EXAMPLE: /part of speech/, /grammatical number/

NOTE 1 The default set of data-categories for TBX were primarily selected from ISO 12620:1999.

NOTE 2 In running text, such as in this International Standard, data-category names are set off using forward slashes and italics. In a TBX document instance, camel case (e.g. <termNote type="partOfSpeech">noun</termNote>) should be used instead of using white space between words.

### 3.7 data-category constraint

specification of the value of an attribute, the content of an element, or one or more structural levels, that constrains the application of a **meta data-category** (3.16)

NOTE The data-category constraints are defined in an XCS file which is used in tandem with a DTD that defines the core-structure module. They can also be included in an integrated schema, such as a Relax NG schema, that incorporates both the core-structure module and the data-category constraints into one file. See also **core-structure module** (3.5).

### 3.8 data file

sequence of bytes that is either stored on a disc in a traditional file system or transmitted as a stream of data over a network

### 3.9 eXtensible Constraint Specification

XCS

XML file that identifies data-categories and their constraints for a specific **TBX TML** (3.21)

### 3.10 extension

totality of objects to which a concept corresponds

[ISO 1087-1:2000]

### 3.11 global information

GI

technical and administrative information applying to the entire data collection

[ISO 16642:2003]

NOTE In a TBX document instance, global information is contained in the front matter.

### 3.12 intension

set of characteristics which makes up the concept

[ISO 1087-1:2000]

### 3.13 interchange

exchange

transaction involving exporting data from one terminological data collection and importing it into another terminological data collection

### 3.14 lemma

lexical unit chosen according to lexicographical conventions to represent the different forms of an inflectional paradigm

[ISO 1951:2007]

### 3.15 lemmatize

to transform an inflected form of a word to its **lemma** (3.13)

### 3.16 lossless roundtrip

series of data manipulation procedures whereby data are output from a database into an interchange format and then re-imported into the same database without loss of information

### 3.17

#### **meta data-category**

name used to group similar **data-categories** (3.6) together; thus, a category of **data-categories** (3.6)

NOTE A meta data-category is equivalent to a typed element in ISO 16642. A meta data-category is instantiated into a terminological data-category through the value of its *type* attribute.

EXAMPLE: In the tag <descrip type="definition">, the meta data-category is *descrip* and the terminological data-category is /definition/.

### 3.18

#### **metadata registry**

information system for registering metadata

NOTE The associated information store or database is known as a metadata register.

### 3.19 object language

language being described in a <langSet>

EXAMPLE In a <langSet xml:lang='fr-FR'> element, the object language is French.

NOTE See also **working language** (3.28).

### 3.20

#### **structural level**

level of the metamodel to which one or more information units can be attached

[ISO 16642:2003]

### 3.21

#### **TBX TML**

**TML** (3.27) that adheres to **TBX** (3.26)

NOTE Implementers of a TBX TML may or may not use an XCS file (in conjunction with a DTD representing the core structure) for validation purposes. Some may choose to use an integrated schema instead.

### 3.22

#### **TBX-default TML**

**TBX** (3.26) and its default selection of **data-categories** (3.6) and their constraints expressed in the default **XCS** (3.9) file

### 3.23

#### **TBX document instance**

file containing terminological entries in a **TBX TML** (3.21) format

### 3.24

#### **terminological database**

database comprising information about special language concepts and terms designated to represent these concepts, along with associated conceptual, term-related, and administrative information

### 3.25

#### **term component**

one of the words comprising a multi-word term, or a component, such as a morpheme, of a single-word term

### 3.26

#### **TermBase eXchange**

TBX

framework consisting of a core structure, and a formalism (**eXtensible Constraint Specification** (3.9)) for identifying a set of **data-categories** (3.6) and their constraints, both expressed in XML

**3.27****terminological markup language**

TML

XML application for describing a terminological data collection conforming to the constraints expressed in ISO 16642 (Terminological markup framework)

NOTE 1 Adapted from ISO 16642:2003.

NOTE 2 TBX coupled with the default XCS file comprise a TML called **TBX-default TML** (3.22). TBX-Basic is also a TML.

**3.28****working language**

default language used in terminological entries

EXAMPLE If definitions, notes, picklist values, and so forth, are normally recorded in English, then English is the working language of the terminological data collection.

NOTE See also **object language** (3.19).

**4 Relationship to other standards**

The foundations for TBX were established by the following three international standards.

- ISO 16642:2003 (TMF) defines the structural metamodel for TBX and other TMLs
- ISO 12620 provides an inventory of data-categories for terminological data
- ISO 12200:1999 (MARTIF) provides the basis for the core structure of TBX and the XML styles of its elements and attributes.

A particular TML requires the choice of an XML style and a selection of data-categories. Most of the data-categories of the TBX-default TML were chosen from ISO 12620:1999, and the XML style of TBX was adopted from ISO 12200. Thus, TBX is a standards-based framework, being based on ISO 16642, ISO 12620, and ISO 12200.

**5 Applications of TBX**

TBX is designed to facilitate the following use cases:

Interchange, such as that required to support

- the flow of terminological data between technologies and systems
- integration of terminological data from multiple sources
- data conversion necessitated by a change in applications or technologies

Dissemination, including

- querying multiple terminological databases through a single user interface by passing data through a common intermediate format on a batch or dynamic basis
- placing data on an online site for download by interested parties
- making entries which require some work available for public feedback
- making terminology available dynamically in networked applications through a Web service

Analysis and representation, including

- comparing the contents of various terminological databases
- studying how lossless a conversion between two terminology databases can be
- designing a new terminological database intended to minimize loss during conversion.

**6 Fundamental principles****6.1 General**

The TBX framework is based on the assumption that, because of the variety of terminological data collections and use scenarios, no one terminological markup language would satisfy all user requirements.

To maximize interoperability, it is recommended that implementers of TBX adhere to ISO standards governing the principles and methodologies of terminology management, and the content and quality of terminological resources, such as those described in [2 Normative references](#) and [Bibliography](#). It is recommended that terminological databases select and use data-categories and their constraints that are specified in this International Standard. Extensions beyond those data-categories and constraints should be taken from ISO 12620 where possible. Fundamental principles of terminological data modelling such as data granularity, data elementarity, data repeatability, and term autonomy, are described in other ISO TC 37 standards.

The information represented in a TBX document instance must be concept-oriented. The terms in a single entry are assumed to be synonymous unless otherwise noted.

TBX allows the representation of various kinds of information about individual terms that distinguish them from other terms in the same concept entry. It also allows for the documentation of directionality in situations where a term in one language may be translated by a given term in another language but the converse is not true due to partial equivalence. It should be noted that some terminological databases document nearly-identical concepts in separate linked entries, while others document nearly-identical concepts in the same entry. TBX can reflect both approaches.

## **6.2 Principles relating to grouping and representing data-categories**

In TBX, there are four general types of data-categories. Understanding what these general types mean and how they are represented will facilitate the understanding of the rest of this International Standard.

NOTE In this specification, attribute names are identified by the *at* sign (@) in Appendix D, and are italicized in running text.

### **core-structure module data-category**

A core-structure module data-category is any data-category that is defined in the core-structure module DTD. For example, `<date>`, `<term>`, and `<descrip>`.

### **meta data-category**

A meta data-category is a core-structure module data-category that takes a *type* attribute, such as `<descrip>`, `<admin>`, and `<termNote>`. It is a *general* data-category that is used for grouping purposes and to reflect the metamodel in ISO 16642. Each *type* attribute value instantiates a meta data-category into a specific terminological data-category that is defined according to ISO 12620. The *type* attribute values are defined in an XCS file. For example, the tag `<descrip type="definition">` comprises the meta data-category `<descrip>` instantiated into a terminological data-category that is called */definition/* according to ISO 12620.

### **data-category implemented using an attribute**

A data-category implemented using an attribute is a terminological data-category that is defined according to ISO 12620, such as */definition/*, and one that is specified as a value of the *name* attribute in the default XCS file. In a TBX document instance, these data-categories appear as the value of a *type* attribute on a meta data-category element. The value of these data-categories is the content of their corresponding element. For instance, the */definition/* data-category, represented via the tag `<descrip type="definition">` takes free text as its content, and the */gender/* data-category, represented in the tag `<termNote type="grammaticalGender">` takes one of a closed set of values (picklist values) as its content (masculine, feminine, neuter, otherGender).

### **data-category implemented as the content of an element**

A data-category implemented as the content of an element is a simple data-category, that is, one value of a closed set of values (picklist). These terminological data-categories are also documented according to ISO 12620. They are enumerated in the default XCS file as the permissible content of a meta data-category having a specific *type* attribute value. For instance, the meta data-category `<termNote>` that has the *type* attribute value 'termType' can have as its content a limited set of values that includes abbreviation, acronym, and so forth. In the integrated RNG schema that is referred to in Annex F, an element's content is constrained to a picklist through embedded Schematron rules. (For a reference to Schematron, see [Bibliography](#).)

The use of meta data-categories in the TBX framework facilitates modularity. The core-structure (which remains

constant) is one module, and a particular XCS file (which expresses constraints on the core structure) is another module. The combination of these two modules defines a particular TML. This approach mirrors TMF (ISO 16642) in that the core-structure module corresponds to the abstract data model of TMF. In addition, it facilitates an explicit description of what two TMLs within the TBX framework have in common (the core structure) and how they differ (expressed as differences between their XCS files). This modular approach is consistent with generally accepted principles of modularity in software engineering, allowing a programmer/analyst to study the core structure and the XCS structure separately without being required to digest multiple large, monolithic schemas.

## 7 Requirements for TBX files

### 7.1 Compliance requirements

For a TML to be compliant with TBX, it shall meet the following three criteria:

1. The TML shall define XML document instances that are valid according to the TBX core-structure module. The core-structure module is described in [8 The core-structure module](#) and is defined formally by the [TBX DTD](#) (Annex A).
2. The TML shall express its data-categories and their constraints in an XCS file that validates against the XCS DTD that is defined in [Annex B \(Normative\) DTD for the data-category constraints \(XCS file\)](#), and it shall adhere to the constraints in that XCS file. A TML that includes a data-category that has the same name as a data-category found in the default XCS shall use this data-category according to its description found in Annex C.
3. The TML can include fewer or more data-categories than those found in the default XCS (Annex C), and still be compliant with TBX, provided that those data-categories are expressed in an XCS file. If the TML includes data-categories that are not in the default XCS, it shall, in addition, describe those additional data-categories in the header of the XCS file.

**NOTE** Several general constraints, such as date formats, are not formally defined in either the DTD or the XCS, but are described in the relevant sections of this document, such as Annex D. These constraints shall also be adhered to for TBX compliance.

The extension for a TBX document instance is *.tbx* and the extension of an XCS document instance is *.xcs*.

Although an XCS document instance must exist to formally define a TBX TML, it need not be used for compliance checking. Indeed, general-purpose XML validation tools do not recognize the constraints in the XCS file unless those constraints have been incorporated into an integrated schema such as the one referred to in Annex F.

TBX compliance checking is schema-definition-language neutral. Three types of compliance checking are described in this document:

#### DTD and XCS

A DTD representing the core structure of TBX is provided at [Annex A \(Normative\) DTD for the core structure module](#). An XCS file representing the default set of data-categories and their constraints is provided in [Annex C.2 XCS file for the default data-categories and constraints](#). With a DTD and an XCS file, a TBX document instance can be validated by using a compliance checker that is specifically designed for TBX files.

#### Relax NG

A Relax NG schema file representing the core structure and the default set of data-categories and their constraints is referred to in Annex F. This file includes some embedded Schematron for some of the data-category constraints. By using this file, one can validate a TBX document instance for compliance with the default TBX TML by using any XML validator that supports Relax NG and Schematron. With appropriate software, an integrated Relax NG schema could be generated for another TBX TML, based on its XCS.

#### Other methods

Compliance checking is also allowed using other methods that incorporate information from the core-structure module and data-category constraints. Additional methods may be documented on the LISA Web.

## ISO 30042:2008(E)

For more information about user-specific TMLs based on TBX, see [12 Creating customized TBX TMLs](#).

### 7.2 Examples of non-compliance

Compliance to TBX includes the following aspects:

1. XML well-formedness
2. validity relative to the core-structure module
3. adherence to the data-category constraints in an XCS file

The following example is not well-formed, since the first `<descrip>` element has a spelling error in the end tag and the `<term>` element has no closing tag.

```
<term>kitten <descrip type='definition'>Small feline</deescrip>
```

The following example is well-formed but not core-structure valid, since the core-structure module of TBX does not allow a `<descrip>` tag to follow a `<tig>`.

```
<langSet xml:lang="fr-ca">
<tig><term>zone de soufflage</term>
</tig>
<descrip type="definition">Area where snow is thrown by a snowplow.</descrip>
</langSet>
```

The following example is valid according to the TBX core-structure DTD but does not adhere to the default XCS, since there is no TBX data-category called "conflagration" in the XCS file.

```
<term>kitten</term> <descrip type='conflagration'>Small feline</descrip>
```

### 7.3 Implementation levels

There are three levels of implementation of TBX for a given software application relative to a particular terminological database:

#### Level 1

The software application shall export and import TBX document instance files that are well-formed and core-structure valid and that adhere to at least one XCS file, and the software application shall detect when document instances are not well-formed or not core-structure valid or not XCS-adherent. The XCS file is not required to be the [default XCS](#), for example, it could be a superset or a subset. Level 1 supports interchange between systems that use the same XCS.

#### Level 2

The software application shall achieve level one implementation but shall also be able to import every data-category that is in the [default XCS](#). Thus, level two implementation supports a degree of blindness in that it can import TBX files from any outside source whose export can be limited to the data-categories that conform to the default XCS file.

#### Level 3

The software application shall achieve level two implementation and in addition be able to check adherence to a comprehensive XCS that supports a lossless roundtrip from the terminological database in the application to a TBX TML and back to the terminological database in the application. Thus, once the information in the terminological database has been exported to TBX, the terminological database can be emptied and subsequently repopulated from the information in the TBX file.

## 8 The core-structure module

### 8.1 Introduction

This section describes the core-structure module for TBX. The elements of the core-structure module are formally declared in Annex A and described in [Annex D](#). For quick access to all these elements, refer to the Index (Annex I).

There is a correspondence between the high-level elements of the core-structure module and the TMF (ISO 16642) metamodel, shown in [Figure 1, High-level structure of the TMF \(ISO 16642\) metamodel](#). The Terminological Data Collection (TDC) corresponds to a TBX document instance.

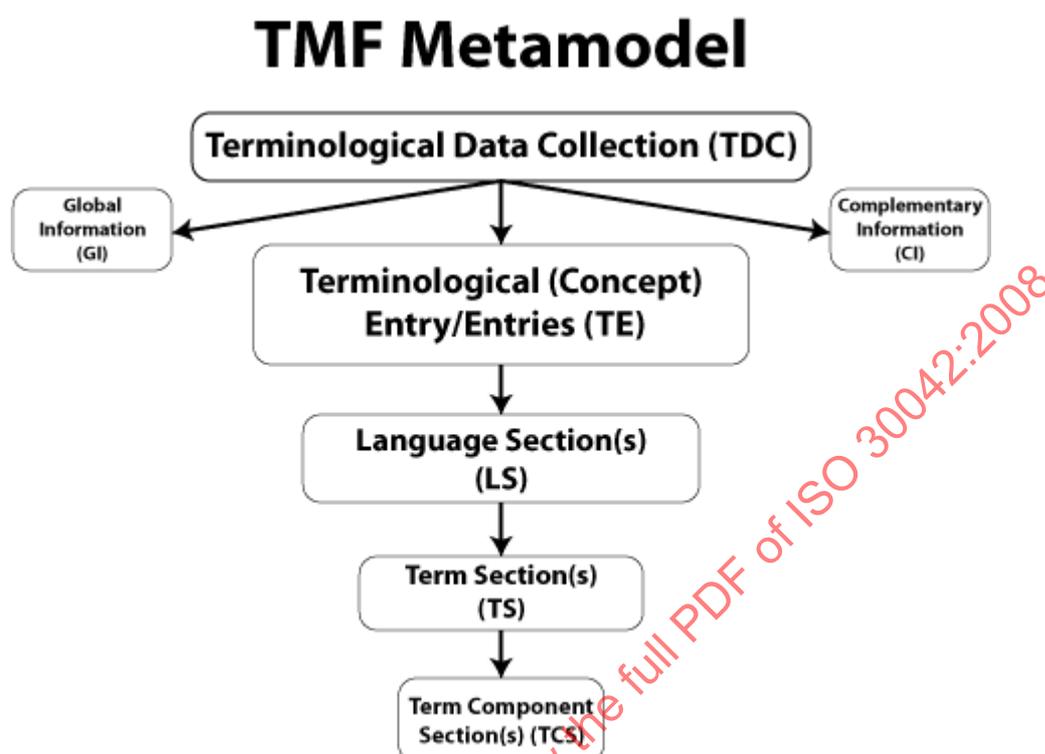


Figure 1. High-level structure of the TMF (ISO 16642) metamodel

In the figures in the following sections, a question mark after an element in the box-and-line diagrams indicates that a single occurrence of the element is optional. A plus sign after an element indicates that one or more occurrences of the element are required. A plus sign in a box by itself indicates that the structure connected to the right of the box occurs one or more times. An asterisk indicates that an element is optional and that multiple instances are allowed. A dotted line indicates a logical exclusive OR relationship, i.e., either of two elements must occur.

## 8.2 Hierarchy

The highest-level XML element in a TBX document instance is the `<martif>` element, which consists of a `<martifHeader>` element and a `<text>` element. (See [Figure 2, The MARTIF structure](#).) These element names are taken from ISO 12200 and have roots in the Text Encoding Initiative.

The `<text>` element in Figure 2 consists of terminological entries, which are enclosed within one `<body>` element, and complementary information (a metamodel object class). In TBX, complementary information is found in the `<back>` element.

The `<martifHeader>` element corresponds to *global information* in the TMF metamodel and consists of a description of the whole terminological data collection (in the `<fileDesc>` element), information about the applicable XCS file (in the `<encodingDesc>` element), and a history of major revisions to the collection (in the `<revisionDesc>` element). Character encoding information shall be included in the header only when the encoding attribute of the XML declaration in the TBX document instance is not a Unicode encoding value.

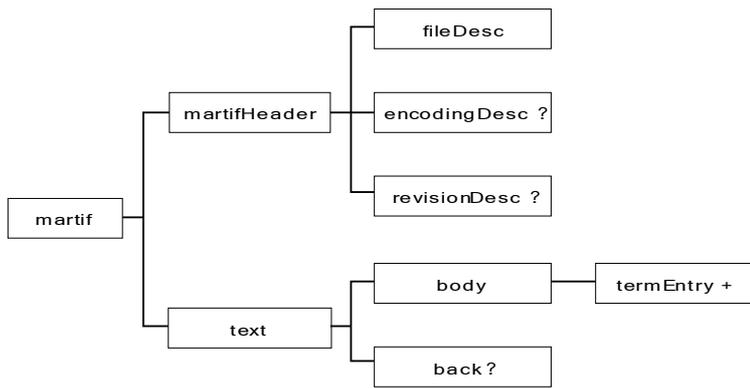


Figure 2. The MARTIF structure

### 8.3 Components of a terminological entry

Each terminological concept entry in the `<body>` element is called a `<termEntry>` and follows the structure of the TMF metamodel.

The *auxInfo* box in [Figure 3. The levels of a terminological entry](#) corresponds to information that can be associated with any one of three levels: the Terminological Entry level (`<termEntry>`, i.e., the concept level), the Language level (`<langSet>`), and the Term level (`<ntig>`, or its simplified version, `<tig>`). The `<termNote>` and `<termNoteGrp>` elements can only appear at the Term level and below. The `<termCompList>` element corresponds to the term component section object class of the TMF metamodel.

A concept entry (termEntry) is hierarchical, as shown in Figure 3.

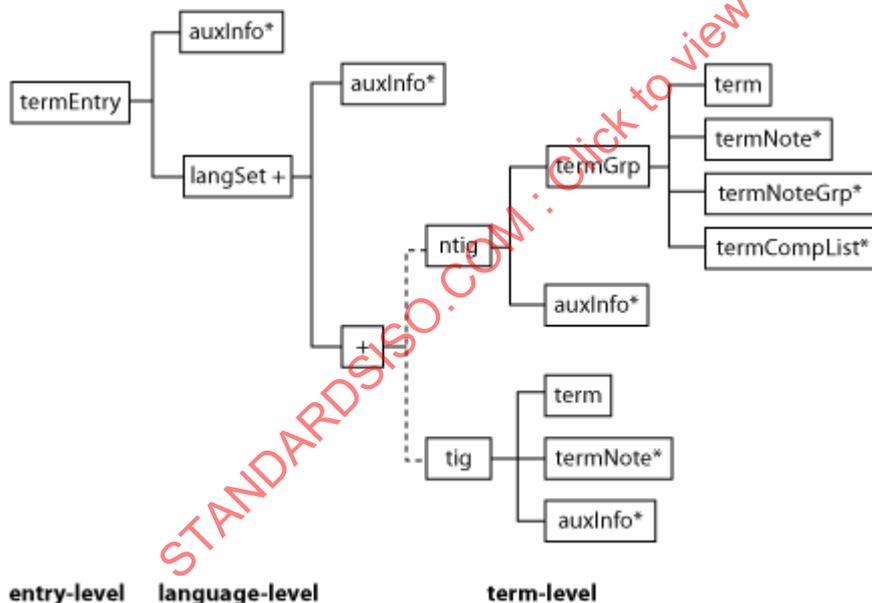


Figure 3. The levels of a terminological entry

### 8.4 Elements that can appear at multiple levels of the entry

The following elements (represented as a group by *auxInfo* in [Figure 3. The levels of a terminological entry](#)), can appear at any of the three levels of an entry.

#### `<descrip>`

Provides descriptive information about the node in question. The type of descriptive information is indicated by the value of the *type* attribute. It can be restricted to certain levels of the entry, depending on the value of the *type* attribute, as indicated in the default XCS file. It can appear alone, or if additional information needs to be associated

with the description, such as administrative information, it can be nested in a [<descripGrp>](#) element.

#### **<descripGrp>**

Contains one [<descrip>](#) element followed by zero or more [<descripNote>](#), [<admin>](#), [<adminGrp>](#), [<transacGrp>](#), [<note>](#), [<ref>](#), and [<xref>](#) elements.

#### **<admin>**

Contains information of an administrative nature for the node in question, such as the source of information, or the project or client for which it applies. The type of administrative information is indicated by the value of the *type* attribute. It can appear alone, or, if additional information needs to be provided, such as a note or a reference, it can be nested in an [<adminGrp>](#) element.

#### **<adminGrp>**

Contains one [<admin>](#) element followed by zero or more [<adminNote>](#), [<note>](#), [<ref>](#), and [<xref>](#) elements.

#### **<adminNote>**

Contains some administrative information, such as the source of a note or other text.

#### **<transacGrp>**

Contains one [<transac>](#) element followed by zero or more [<transacNote>](#), [<date>](#), [<note>](#), [<ref>](#), and [<xref>](#) elements. It encloses information about a transaction, such as the date it was carried out, or the person who performed it. A date is specified by a [<date>](#) element, and the name of a person or entity responsible for the transaction is specified by a [<transacNote>](#) element.

NOTE Any date in a TBX entry shall appear within a [<transacGrp>](#) element.

#### **<note>**

Contains any kind of note.

#### **<ref>**

A cross-reference that points to another element within the [<marfil>](#) element.

#### **<xref>**

A cross-reference that points to an external object using a URI (a URL or other Web address).

## **8.5 Elements that occur only at the term level or lower**

The term level in the entry hierarchy is referred to by *ntig* and *tig* in [Figure 3, The levels of a terminological entry](#). The topmost element at the term level can be either [<ntig>](#) or its simpler counterpart [<tig>](#).

The following elements can occur at the term level.

#### **<tig>**

Provides all the information about a term. It contains one [<term>](#) element followed by zero or more [<termNote>](#) elements followed by any of the elements described in [8.4 Elements that can appear at multiple levels of the entry](#).

#### **<ntig>**

Provides all the information about a term. It supports a deeper nesting structure than its simpler counterpart [<tig>](#). It contains one [<termGrp>](#) element followed by any of the elements described in [8.4 Elements that can appear at multiple levels of the entry](#). This element is chosen, instead of [<tig>](#), if you want to document the components of a term within the same entry.

#### **<termGrp>**

Contains one [<term>](#) element, followed optionally by either a [<termNote>](#) or a [<termNoteGrp>](#) element, and zero or more [<termCompList>](#) elements.

#### **<termNoteGrp>**

Contains one [<termNote>](#) element and zero or more [<admin>](#), [<adminGrp>](#), [<transacGrp>](#), [<note>](#), [<ref>](#), and [<xref>](#) elements.

#### **<termNote>**

Provides different types of information about a term, depending on the value of its *type* attribute. For instance, it can

contain grammatical information such as gender or part of speech, references to other related terms, administrative information such as the source of the term or what project it is used for, and so forth.

The following elements can occur at the term component level.

### <termCompGrp>

Contains one <termComp> element followed by, optionally, either a <termNote> or a <termNoteGrp> element, and zero or more <admin>, <adminGrp>, <transacGrp>, <note>, <ref>, and <xref> elements that apply to the term component.

### <termCompList>

Shows the internal composition of a term and consists of any of the elements described in [8.4 Elements that can appear at multiple levels of the entry](#) followed by one or more <termComp> or <termCompGrp> elements.

### <termComp>

Contains some component of a term, such as one of the words of which it is composed. It can be contained in a <termCompGrp> or a <termCompList>.

## 8.6 Handling of text

### 8.6.1 Types of text

In TBX there are five different ways in which an element may contain text.

#### noteText

*noteText* is an entity defined in the core DTD. Elements that are defined with this entity can take text as their content, and the text itself may contain some pre-defined inline markup elements. *noteText* is used by textual elements that may require some inline markup, such as <descrip type="definition"> or <termNote type="usageNote">.

#### basicText

*basicText* is also an entity defined in the core DTD. Elements that are defined with this entity can take text as their content, plus one or more <hi> elements to represent more restricted inline markup styles than that permitted by *noteText*. *basicText* is used by <term> and term-like elements such as <termNote type="shortFormFor">.

#### plainText

*plainText* refers to text without any inline markup. This corresponds to the Relax NG definition <rng:text/> and to the DTD definition PCDATA. Elements that use this data type are those whose content can be considered a token, such as <descrip type="subjectField"> and <admin type="productSubset">.

#### text as markup

Some elements, such as <bpt>, contain text that actually represents markup from another environment, such as a translation memory system, that needs to be retained in TBX. These elements are described in [8.6.2 Inline markup elements](#).

#### picklist

Some elements take a further constrained implementation of *plainText*, in the form of a pre-defined set of picklist values. For example, the element <termNote type="grammaticalGender"> can only contain one of four *plainText* strings according to the default XCS file: masculine, feminine, neuter, and otherGender. The picklist values are specified in the XCS file.

The data type supported by each TBX data-category is specified in [9 The default data-category constraints](#).

### 8.6.2 Inline markup elements

TBX allows five elements for inline markup within data-categories having a data type of *noteText* as defined in [9.2 Data-categories built into the core structure DTD of TBX](#). These elements are used to specify the language or function of textual strings or to allow TBX document instances to contain various kinds of other markup (such as HTML or text-processing markup) that needs to be retained but should not necessarily be processed during terminology management functions. These elements are described below:

**<foreign>**

Foreign - The **<foreign>** element is used to mark a segment of text that is in a different language from the surrounding text. The optional *xml:lang* attribute may be used to identify the language of the text contained in this element.

**<hi>**

Highlight - The **<hi>** element delimits a section of text. It can be used for various processing tasks. For example, it can be used to mark a mathematical expression, or an entailed term in a definition or other text field. The **<hi>** element is also allowed within **<term>** elements to mark portions of terms. For example, there are terms that have a character in subscript or superscript, or a component within the term that is italicized or requires other formatting.

**<bpt>**

Begin paired tag - The **<bpt>** element is used to delimit the first of a paired sequence of native codes (for example, an HTML **<strong>** tag). Each **<bpt>** shall have a corresponding **<ept>** element after it within the same parent element.

**<ept>**

End paired tag - The **<ept>** element is used to delimit the second of a paired sequence of native codes (for example, an HTML **</strong>** tag). Each **<ept>** shall have a corresponding **<bpt>** element before it within the same parent element.

**<ph>**

Placeholder - The **<ph>** element is used to delimit a sequence of native standalone codes in the segment that is not paired with another element (e.g., an XHTML **<br />** code). It should not be used to replace codes that occur in logical pairs.

Any enclosed markup shall have any start-tag characters ("**<**") or ampersands (**&**) converted into their respective entities, **&lt;** and **&amp;**. This conversion allows the markup to be retained and processed during display or import without complicating the TBX core structure by having to include the XHTML DTD in the core structure. Any kind of text-based markup, including RTF, can be encapsulated in these tags and later retrieved without loss of information.

The *i* (internal matching) attribute is used to pair **<bpt>** elements with their corresponding **<ept>** elements. This mechanism provides TBX with support to mark up possibly overlapping ranges of codes, as in the following incorrect, but nevertheless common HTML construction:

This is `<i>some <b>sample HTML</i> markup </b>` with improper nesting.

The above markup would be represented as follows using **<ept>** and **<bpt>**:

This is `<bpt i="1">&lt;i>some <bpt i="2">&lt;b>sample HTML<ept i="1">&lt;/i></ept> markup <ept i="2">&lt;/b></ept>` with improper nesting.

Values of *i* shall be unique within the parent element of the **<bpt>** or **<ept>** element.

### 8.6.3 xml:lang attribute

Every element containing free text shall have the language of that text indicated by having an explicit or inherited *xml:lang* attribute. In TBX, the *xml:lang* attribute does not apply to any attributes or their values.

### 8.6.4 Encoding

In TBX, all text shall be in Unicode (ISO/IEC 10646). There are three allowable encodings of Unicode: UTF-8, UTF-16, and seven-bit ASCII with non-ASCII characters represented as hex character references to their Unicode code point. Hex character references shall be represented as specified in the XML standard.

### 8.6.5 Mathematical expressions in TBX

At times it is necessary to represent mathematical expressions in TBX document instances, including within terms and definitions. For mathematical expressions that require formatting beyond that available in a simple text format, when validation against the TBX core DTD and an XCS file is required, the **<hi>** element with a *type* attribute value of 'math' may be used. As the DTD mechanism does not provide support for the XML namespace mechanism, the content of this element should consist of LaTeX-format data (which does not use XML-style tagging). The following

examples show how LaTeX mathematical expressions may be incorporated into a TBX document instance. (For more information on LaTeX mathematical formatting, consult the TeX Users' Group at [www.tug.org](http://www.tug.org).) Note that any instances of < or & in LaTeX strings shall be replaced with the corresponding entities: &lt; and &amp; respectively.

Mathematical example #1:

```
<descrip type="definition">
result determined by the expression <hi type="math"> |C| = \sqrt{C'^{2} + C''^{2}}
= \sigma_{2} / \epsilon_{2} </hi>
</descrip>
```

The graphical equivalent of the contents of the <hi> element in this example is the following:

$$|C| = \sqrt{C'^2 + C''^2} = \sigma_2 / \epsilon_2$$

Mathematical example #2:

```
<descrip type="definition">
ratio of actual to critical damping, where critical damping is that required for the
borderline condition between oscillatory and non-oscillatory behaviour. Damping ration
is a function of the logarithmic decrement L: It is dimensionless. Equation: <hi
type="math"> \mu = \frac{A/2\pi}{\sqrt{1+(a/2\pi)^2}} = \sin \arctan (A/2\pi)
</hi>For small values of<hi type="math">A/\mu = A/2\pi</hi>
</descrip>
```

The two instances of <hi> may be graphically represented as follows:

$$\mu = \frac{A/2\pi}{\sqrt{1+(a/2\pi)^2}} = \sin \arctan (A/2\pi)$$

and

$$A/\mu = A/2\pi$$

If desired, the MathML namespace may be used instead of LaTeX markup. However, if you are using the core structure DTD for validation, some preprocessing will be required, namely, commenting out or removing the MathML namespace elements. The validation of MathML is not possible with the core DTD. No such processing is required when validating against a schema that supports XML namespaces.

## 8.7 Meta data elements

A meta data element is an element that can have multiple different content models, one for each different value of its *type* attribute. By keeping the element name constant while further specifying the element at the attribute level, a meta data element serves the purpose of grouping together data-categories that have common properties. For instance, the <descrip> element contains various descriptions, and the <termNote> element contains various information about terms. The content constraints of meta data elements are defined in the [XCS file](#), or in an integrated schema such as the one in Annex F. The combination of the element name and the value of its *type* attribute corresponds to a data-category as defined in ISO 12620. For instance, the element <descrip type="definition"> corresponds to the /definition/ data-category from ISO 12620.

The following are the meta data elements in TBX:

- <termNote>
- <termCompList>
- <admin>
- <adminNote>
- <transac>
- <transacNote>
- <descrip>
- <descripNote>

- [<ref>](#)
- [<xref>](#)
- [<refObjectList>](#)

There are two cases where the data-category is not identified by the value of the *type* attribute on the meta data element which it logically characterizes. The first case involves the *list* data elements, i.e. [<termCompList>](#) and [<refObjectList>](#). These two data elements take a *type* attribute but the meaning of this *type* attribute actually applies to the components of the list, not to the list itself. For instance, for [<termCompList>](#) the permissible *type* attributes include 'hyphenation', 'morphological element', etc. These characteristics logically apply to the term components in the list, not to the list itself. This is an example of the principle of inheritance, where the child elements inherit the *type* property of their parent.

The second case involves the *group* data elements, such as [<transacGrp>](#) and [<descripGrp>](#). None of these elements are allowed to have a *type* attribute. This is to permit both flat structures and group structures for the data elements in question. Instead, the *type* of the group is indicated in the content of its principal child element. See the example of a *transacGrp* in 10.3.4. It shows that the *type* of transaction which the [<transacGrp>](#) describes is indicated as the content of the child element [<transac>](#). This is an example of the principle of reverse inheritance.

Unlike the other meta data elements, the elements [<adminNote>](#) and [<transac>](#) are instantiated as only one data-category because they only have one permissible value of their *type* attribute, i.e. [<adminNote>](#) only takes the *type* attribute 'sourceType', and [<transac>](#) only takes the *type* attribute 'transactionType'. However, they are still considered meta data elements because for extensibility purposes new *type* attribute values could be defined by a user group in a custom XCS file.

## 8.8 Attributes

The main attributes used in TBX are *xml:lang* (language), *type* (to identify the data-category), *id* (to identify an element uniquely within the XML document), and *target* (to point to another element or an external object). The value of the *xml:lang* attribute inherits downward through the implied tree structure of the XML document unless overridden by another *xml:lang* attribute. The [<martif>](#) element is required to have an *xml:lang* attribute. The language specified in the [<martif>](#) element becomes the working language of the entire TBX file. Each [<langSet>](#) element shall also specify a language that applies to that language section. This becomes the object language and the working language of the language section. Thus, the content of a [<descrip type="definition">](#) element at the entry level is assumed to be in the working language of the TBX file unless otherwise specified, and a note in a language section is assumed to be in the same language as the language section unless otherwise specified by an explicit *xml:lang* attribute. The allowed values of the *xml:lang* attribute in TBX are found in IETF RFC 4646 or its successor, as identified in IETF BCP 47.

The *id* and *target* attributes work together to point unambiguously between elements in the same TBX file. For example, one entry:

```
<termEntry id="eid-database-5574">
... (entry for "hunting dog")
</termEntry>
```

could be pointed to by another entry:

```
<termEntry>
<descrip type="superordinateConceptGeneric" target="eid-database-5574">hunting dog
</descrip>
... (entry for "Retriever" [a type of hunting dog])
</termEntry>
```

The content "hunting dog" in the second entry is for display purposes. It provides a visible label for the link to the target entry.

The *target* attribute points to an external object when the value of the corresponding *type* attribute begins with the letter x.

## 8.9 Character set issues

A variety of mnemonic entities are allowed in ISO 12200, but TBX is more restrictive. Only hex character references, and the mnemonic entities built into XML, are used. This is to reduce the burden on a blind import routine, which cannot anticipate all the mnemonic character entities that might be used. TBX data files shall be in one of three encodings of Unicode: (a) UTF-16 (b) UTF-8 or (c) pure 7-bit ASCII (ISO/IEC 646) in which non-ASCII characters are encoded as eight ASCII characters using an XML hex character reference (such as `&#1103;`) as well as the five pre-defined XML entities: quotation mark, ampersand, apostrophe, less-than sign and greater-than sign.

Such hex character references are automatically converted to readable characters (in this case: я) when an XML data file containing them is displayed in various types of software that can display XML files, such as many web browsers. This third type of encoding can formally be considered UTF-8, although it does not use the UTF-8 method of encoding characters whose code point is above 127.

## 8.10 Language

The `<langSet>` and `<lang>` elements are the only elements for which the `xml:lang` attribute is mandatory.

On the `<lang>` element, this attribute indicates the working language of the TBX document instance. It applies to all elements in the TBX document instance unless overridden by another `xml:lang` attribute, such as in a `<langSet>` element.

On the `<langSet>` element, this attribute indicates both the working language and the object language of the corresponding language section. Unless overridden by another `xml:lang` attribute on a child element, the content of all elements in a language section are assumed to be in the language specified by the `xml:lang` attribute of that `<langSet>` element. This does not apply to the content of elements that are picklist values as defined in the XCS file. To change the language of an element in a language section to a value other than the value of the `xml:lang` attribute on the `<langSet>` element, use an `xml:lang` attribute on that element.

# 9 The default data-category constraints

## 9.1 Introduction

This section describes the default TBX data-categories and their constraints. These data-categories and constraints are formally described in [Annex D](#). A TBX document instance can be validated against the default data-categories and constraints by using the XCS file that is provided in [Annex C](#).

The default XCS is not the only XCS that can be used for TBX-compliant terminological databases. Particular user groups can define their own XCS file, as described in [12 Creating customized TBX TMLs](#).

Guidelines for encoding data-categories as XML elements are given in [10.2 Examples of encoding TBX elements](#).

The data-categories in the following tables are organized in groups that reflect common properties or purpose. Each data-category, other than the basic data-categories built into the core DTD in the first section below, is related to the metamodel by being classified as either administrative or descriptive. Descriptive data-categories describe either a concept or a term. With a few exceptions, data-categories that use the element `<descrip>` describe the concept, and data-categories that use the element `<termNote>` describe the term. All data-categories that use the element `<admin>` are administrative. A data-category that is allowed a `target` attribute (as indicated in the default XCS file) indicates a relation to another piece of information, such as another term or concept, or a bibliographic reference.

The data-categories in section 9.2 are implemented directly as TBX elements or attributes in the core-structure DTD. The remaining tables contain data-categories that are specializations (through values of the `type` attribute) of the meta data-categories, which themselves are defined in the core-structure DTD. The first column is the unique name of the data-category (from ISO 12620). The second column (Data type) indicates what kind of text is allowed in the element. The third column indicates whether this element can take a `target` attribute, in which case it indicates what kind of element can be targeted. The fourth column indicates which meta data-category is used for this data-category. The fifth column indicates the levels in the metamodel at which a particular data-category can appear.

If the data type indicated in the tables is *picklist*, the content of the element is restricted to a pre-defined set of

values, sometimes called a *value domain*. Picklist values are specified in the link in the corresponding *Meta data-category* column, where they are listed as permissible values of the content of the element. The picklist values are also defined in the default XCS file. User groups can choose their own display values for these picklists, as described in [12.3 Creating customized picklist display names](#).

**NOTE** The meanings of the picklist values themselves are not defined in this International Standard. To obtain information about the meaning of picklist values, refer to ISO 12620.

By default, the `<admin>` elements can appear at any level, and `<descrip>` elements can appear at the entry, language, or term levels unless restricted to a specific level in the default XCS file. In the XCS file, the levels are expressed using the `<levels>` element with content values termEntry (term entry level), langSet (language section level), or term (term level). A fourth possible level, for individual components of a term, is expressed with a `<forTermCompGrp>` element. The element `<termNote>` can appear only at the term level, unless authorized (by a `<forTermCompGrp>` element) to appear at the term component level as well.

Data-categories that are not constrained to a picklist in the TBX default XCS can be constrained to a picklist in a user-group subset XCS (see [12.1 General information about TMLs](#)). One obvious candidate for a user-group picklist is partOfSpeech, which for many languages can have standardized values.

## 9.2 Data-categories built into the core structure DTD of TBX

### 9.2.1 Elements

Data-category name	Data type	Target	Element
begin paired tag	plainText	none	<code>&lt;bpt&gt;</code>
date	date (ISO format)	none	<code>&lt;date&gt;</code>
end paired tag	plainText	none	<code>&lt;ept&gt;</code>
foreign	noteText	none	<code>&lt;foreign&gt;</code>
highlight	plainText	element	<code>&lt;hi&gt;</code>
note	noteText	none	<code>&lt;note&gt;</code>
placeholder	plainText	none	<code>&lt;ph&gt;</code>
term	basicText	none	<code>&lt;term&gt;</code>

### 9.2.2 Attributes

Data-category name	Data type	Attribute
elementIdentifier	CDATA	id
lang	language codes	xml:lang
link	CDATA, IDREF	target

## 9.3 Data-categories specialized from meta data-categories through the default XCS file

### 9.3.1 Properties and descriptions of concepts

These data-categories describe properties of concepts. Some are allowed at the term level.

Data-category name	Data type	Target	Meta data-category	Level
audio	plainText	binaryData	<code>&lt;descrip&gt;</code>	langSet, termEntry, term
characteristic	plainText	none	<code>&lt;descrip&gt;</code>	term
conceptOrigin	plainText	none	<code>&lt;admin&gt;</code>	langset, termEntry, term
context	noteText	none	<code>&lt;descrip&gt;</code>	term
contextType	picklist	none	<code>&lt;descripNote&gt;</code>	langset, termEntry, term
definition	noteText	none	<code>&lt;descrip&gt;</code>	langSet, termEntry, term

Data-category name	Data type	Target	Meta data-category	Level
definitionType	picklist	element	<descripNote>	langset, termEntry, term
example	noteText	none	<descrip>	langSet, termEntry, term
explanation	noteText	none	<descrip>	langSet, termEntry, term
figure	plainText	binaryData	<descrip>	langSet, termEntry, term
otherBinaryData	plainText	binaryData	<descrip>	langSet, termEntry, term
quantity	plainText	none	<descrip>	term
range	plainText	none	<descrip>	term
sampleSentence	noteText	none	<descrip>	term
table	plainText	binaryData	<descrip>	langSet, termEntry, term
unit	plainText	none	<descrip>	term
video	plainText	binaryData	<descrip>	langSet, termEntry, term

### 9.3.2 Concept relations

These data-categories indicate relations between concepts.

Data-category name	Data type	Target	Meta data-category	Level
antonymConcept	basicText	entry	<descrip>	termEntry
associatedConcept	basicText	entry	<descrip>	langSet, termEntry
broaderConceptGeneric	basicText	entry	<descrip>	langSet, termEntry
broaderConceptPartitive	basicText	entry	<descrip>	langSet, termEntry
conceptPosition	plainText	conceptSysDescrip	<descrip>	langSet, termEntry
coordinateConceptGeneric	basicText	entry	<descrip>	langSet, termEntry
coordinateConceptPartitive	basicText	entry	<descrip>	langSet, termEntry
relatedConcept	basicText	entry	<descrip>	langSet, termEntry
relatedConceptBroader	basicText	entry	<descrip>	langSet, termEntry
relatedConceptNarrower	basicText	entry	<descrip>	langSet, termEntry
sequentiallyRelatedConcept	basicText	entry	<descrip>	langSet, termEntry
spatiallyRelatedConcept	basicText	entry	<descrip>	langSet, termEntry
subordinateConceptGeneric	basicText	entry	<descrip>	langSet, termEntry
subordinateConceptPartitive	basicText	entry	<descrip>	langSet, termEntry
superordinateConceptGeneric	basicText	entry	<descrip>	langSet, termEntry
superordinateConceptPartitive	basicText	entry	<descrip>	langSet, termEntry
temporallyRelatedConcept	basicText	entry	<descrip>	langSet, termEntry

### 9.3.3 Properties and descriptions of terms

This group of data-categories include various grammatical, usage, and administrative properties of terms.

Data-category name	Data type	Target	Meta data-category	Level
animacy	picklist	none	<termNote>	term, termComponent
etymology	noteText	none	<termNote>	term, termComponent
frequency	picklist	none	<termNote>	term
geographicalUsage	plainText	none	<termNote>	term
grammaticalGender	picklist	none	<termNote>	term, termComponent
grammaticalNumber	picklist	none	<termNote>	term, termComponent
grammaticalValency	plainText	none	<termNote>	term
language-planningQualifier	picklist	none	<termNote>	term
normativeAuthorization	picklist	none	<termNote>	term
partOfSpeech	plainText	none	<termNote>	term, termComponent
proprietaryRestriction	picklist	none	<termNote>	term

Data-category name	Data type	Target	Meta data-category	Level
register	picklist	none	<termNote>	term
reliabilityCode	picklist	none	<descrip>	langSet, termEntry, term
temporalQualifier	picklist	none	<termNote>	term
termLocation	plainText	none	<termNote>	term
termProvenance	picklist	none	<termNote>	term
timeRestriction	plainText	none	<termNote>	term
transferComment	noteText	term	<termNote>	term
usageNote	noteText	none	<termNote>	term

### 9.3.4 Types of terms, relations to terms

This group includes data-categories that describe the different types of terms, or indicate relations to other terms.

Data-category name	Data type	Target	Meta data-category	Level
abbreviatedFormFor	basicText	term	<termNote>	term
antonymTerm	basicText	term	<termNote>	term
directionality	picklist	term	<termNote>	term
falseFriend	basicText	term	<termNote>	term
homograph	basicText	term	<termNote>	term
shortFormFor	basicText	term	<termNote>	term
termType	picklist	none	<termNote>	term

### 9.3.5 Components of terms

Data-category name	Data type	Target	Meta data-category	Level
hyphenation	plainText	none	<termCompList>	termComponent
lemma	plainText	none	<termCompList>	termComponent
lionHotkey	plainText	none	<termNote>	term
morphologicalElement	plainText	none	<termCompList>	termComponent
pronunciation	basicText	none	<termNote>	term, termComponent
syllabification	plainText	none	<termCompList>	termComponent
termElement	plainText	none	<termCompList>	termComponent
termStructure	plainText	none	<termNote>	term, termComponent

### 9.3.6 Subsetting

This table includes data-categories that are used to divide the database into logical subsets such as the subject field and various administrative subsets.

Data-category name	Data type	Target	Meta data-category	Level
applicationSubset	plainText	none	<admin>	langset, termEntry, term
businessUnitSubset	plainText	none	<admin>	langset, termEntry, term
customerSubset	plainText	none	<admin>	langset, termEntry, term
environmentSubset	plainText	none	<admin>	langset, termEntry, term
productSubset	plainText	none	<admin>	langset, termEntry, term
projectSubset	plainText	none	<admin>	langset, termEntry, term
securitySubset	picklist	none	<admin>	langset, termEntry, term
subjectField	plainText	none	<descrip>	termEntry
subsetOwner	plainText	none	<admin>	langset, termEntry, term

### 9.3.7 Search, indexing, other usages of terms

Data-category name	Data type	Target	Meta data-category	Level
classificationCode	plainText	bibl	<descrip>	langSet, termEntry, term
indexHeading	plainText	none	<admin>	langset, termEntry, term
keyword	plainText	none	<admin>	langset, termEntry, term
searchTerm	plainText	none	<admin>	langset, termEntry, term
sortKey	plainText	none	<admin>	langset, termEntry, term
thesaurusDescriptor	plainText	thesaurusDescrip	<descrip>	termEntry

### 9.3.8 Administrative information

Data-category name	Data type	Target	Meta data-category	Level
administrativeStatus	picklist	none	<termNote>	term
annotatedNote	noteText	none	<admin>	langset, termEntry, term
databaseType	plainText	none	<admin>	langset, termEntry, term
domainExpert	plainText	bibl	<admin>	langset, termEntry, term
elementWorkingStatus	picklist	none	<admin>	langset, termEntry, term
entrySource	plainText	none	<admin>	langset, termEntry, term
noteSource	plainText	none	<adminNote>	langset, termEntry, term
originatingDatabase	plainText	none	<admin>	langset, termEntry, term
originatingInstitution	plainText	none	<admin>	langset, termEntry, term
originatingPerson	plainText	none	<admin>	langset, termEntry, term
processStatus	picklist	none	<termNote>	term
responsibility	plainText	respPerson	<transacNote>	langset, termEntry, term
source	noteText	none	<admin>	langset, termEntry, term
sourceIdentifier	plainText	bibl	<admin>	langset, termEntry, term
sourceType	picklist	none	<adminNote>	langset, termEntry, term
transactionType	picklist	none	<transac>	langset, termEntry, term
usageCount	plainText	none	<transacNote>	langset, termEntry, term

### 9.3.9 References and cross-references

Data-category name	Data type	Target	Meta data-category	Level
corpusTrace	plainText	external	<xref>	langset, termEntry, term
crossReference	plainText	element	<ref>	langset, termEntry, term
externalCrossReference	plainText	external	<xref>	langset, termEntry, term
see	plainText	element	<ref>	langset, termEntry, term, termComponent
xAudio	plainText	external	<xref>	langset, termEntry, term
xGraphic	plainText	external	<xref>	langset, termEntry, term
xMathML	plainText	external	<xref>	langset, termEntry, term
xSource	plainText	external	<xref>	langset, termEntry, term
xVideo	plainText	external	<xref>	langset, termEntry, term

### 9.3.10 Inline markup

This group of data-categories provide the means to mark up sections of text in text fields such as the definition.

Data-category name	Data type	Target	Meta data-category	Level
bold	plainText	none	<hi>	langset, termEntry, term
entailedTerm	plainText	none	<hi>	langset, termEntry, term

hotkey	plainText	none	<hj>	langset, termEntry, term
italics	plainText	none	<hj>	langset, termEntry, term
math	plainText	none	<hj>	langset, termEntry, term
subscript	plainText	none	<hj>	langset, termEntry, term
superscript	plainText	none	<hj>	langset, termEntry, term

## 10 Examples

### 10.1 Example of a typical TBX file

The following is an example of a simple but complete TBX file. The components are described after the example.

```
<?xml version='1.0'?> <!DOCTYPE martif SYSTEM "TBXcoreStructV02.dtd">
<martif type="TBX" xml:lang="en">
  <martifHeader>
    <fileDesc>
      <sourceDesc>
        <p>From an Oracle corporation termbase</p>
      </sourceDesc>
    </fileDesc>
    <encodingDesc>
      <p type="XCSURI">http://www.lisa.org/fileadmin/standards/tbx/TBXXCSV02.XCS</p>
    </encodingDesc>
  </martifHeader>
  <text>
    <body>
      <termEntry id="eid-Oracle-67">
        <descrip type="subjectField">manufacturing</descrip>
        <descrip type="definition">A value between 0 and 1 used in ...</descrip>
        <langSet xml:lang="en">
          <tig>
            <term id="tid-Oracle-67-en1">alpha smoothing factor</term>
            <termNote type="partOfSpeech">noun</termNote>
          </tig>
        </langSet>
        <langSet xml:lang="hu">
          <tig>
            <term id="tid-Oracle-67-hu1">Alfa simitási tényező</term>
            <termNote type="partOfSpeech">noun</termNote>
          </tig>
        </langSet>
      </termEntry>
    </body>
  </text>
</martif>
```

The `xml` declaration states that the following lines constitute an XML document that conforms to version 1.0 of the definition of XML by the World Wide Web consortium (W3C). The `DOCTYPE martif...` declaration states that this particular XML document can be validated against a specification of the TBX core structure, which, for this document, is an XML DTD called `TBXcoreStructV02.dtd`. (Alternatively, the document structure can be validated against a schema version of the description of the core structure, such as the one referred to in Annex F.)

The `<martif>` element indicates, with its `type` attribute, that it is a TBX document. The `xml:lang` attribute indicates that the default language for text in this document is English (ISO 639 code 'en'). The `xml:lang` attribute can take an ISO 639 code as its value, but can also take a two-part value, such as *fr-CA* for Canadian French, with the second component taken from ISO 3166 (country codes). Further augmentations can include script codes from ISO 15924, dialects, and other special features introduced in IETF RFC 4646 or its successor, as defined in IETF BCP 47.

The `<martifHeader>` contains other elements that provide global information about the collection: specifically, a file

description indicating that the example was derived from an entry in a terminological database used at Oracle corporation and that the TBX XCS (TBXXCSV02.XCS) contains the additional data-category constraints.

The `<text>` element surrounds the `<body>` element. The purpose of `<text>` element is to maintain compatibility with the Text Encoding Initiative guidelines. The `<body>` element contains the collection of concept-oriented *Terminological Entry* (`<termEntry>`) elements, and, optionally, the `<back>` element.

Each `<termEntry>` element is one instance of the *Terminological Entry* object class in ISO 16642. The *id* attribute has a value that is unique throughout the document, making it possible for other elements to point unambiguously to this element. The id `eid-Oracle-67` consists of the information: `eid` [entry identifier] + the name of the database [Oracle] + the serial number of the entry (67).

The `<descrip type='subjectField'>` element specifies a subject field for the entry. The */subject field/* data-category is authorized by the XCS file. It consists of a meta data-category element (`<descrip>`) with the specific data-category indicated in the value of the *type* attribute.

The `<descrip type='definition'>` element contains a definition of the concept.

The `<langSet>` element corresponds to a *Language Section* object class, according to which a terminological entry consists of associated information and language sections.

The metamodel in ISO 16642 states that a Language Section consists of instances of a *Term Section* object class, which, in TBX corresponds to a `<tig>` (or `<ntig>`) element. An instance of a Term Section consists of a term and associated information, which in this case is the type of term, expressed as a *termType*. The name *tig* stands for *term information group*. The id `tid-Oracle-67-en1` consists of the information: `tid` [term identifier] + the name of the database [Oracle] + the serial number of the entry [67] + the language code [en] + the serial number of the `<tig>` within that language group [1].

The `<termNote>` element with the *termType* attribute corresponds to the data-category */term type/*. Its value in this case is 'fullForm'. A `<termNote>` tag is used since the information is associated with the term rather than with the concept.

The second `<langSet>` element, having the language attribute value 'hu', begins the Hungarian language section. The second `<tig>` element group consists of a term section with a Hungarian term but no definition and no explicit term type.

This sample TBX entry has several properties:

- It corresponds directly to the metamodel in ISO 16642.
- It is well-formed and core-structure-valid.
- It adheres to the default TBX extensible constraint specification (XCS).

## 10.2 Examples of encoding TBX elements

### 10.2.1 Element with one attribute

Many TBX elements require only one attribute, such as the *type* attribute. They would appear as follows:

```
<descrip type="definition">Content of the element goes here, in this case, a definition.</descrip>
```

### 10.2.2 Elements with two or more attributes

Some TBX elements include two attributes, such as all those elements that allow a *target* attribute. The following markup might be used in an entry about the term ACL.

```
<termNote type="abbreviatedFormFor" target="IBM-05529">access control list</termNote>
```

Elements occasionally have three or even more attributes, such as to specify the language and an identifier of an element, as in the following example, which might occur if the abbreviation *XML* is used in languages other than English, requiring the following markup in the non-English language section, if there is a need to indicate the language of the element content:

```
<termNote type="abbreviatedForm" target="x-5966" xml:lang="en-US">XML</termNote>
```

## 10.3 Examples of TBX entries

### 10.3.1 Term components

The following sample demonstrates the description of term components. In addition it also shows the possibility of using a <tig> in place of an <ntig> in simple cases:

- fr: table des transitions d'états
- en: state transition table

The following <langSet> element contains an <ntig> followed by a <tig>:

```
<langSet xml:lang="fr">
  <ntig>
    <termGrp>
      <term>table des transitions d'états</term>
      <termCompList type="termElement">
        <termCompGrp>
          <termComp>table</termComp>
          <termNote type="grammaticalGender">feminine</termNote>
        </termCompGrp>
        <termCompGrp>
          <termComp>des</termComp>
          <termNote type="partOfSpeech">other</termNote>
        </termCompGrp>
        <termCompGrp>
          <termComp>transitions</termComp>
          <termNote type="grammaticalNumber">plural</termNote>
          <termNote type="grammaticalGender">feminine</termNote>
        </termCompGrp>
        <termCompGrp>
          <termComp>de</termComp>
          <termNote type="partOfSpeech">preposition</termNote>
        </termCompGrp>
        <termComp>états</termComp>
      </termCompList>
    </termGrp>
  </ntig>
</langSet>
<langSet xml:lang="en">
  <tig>
    <term>state transition table</term>
  </tig>
</langSet>
```

The following ntig is equivalent to the English tig given above:

```
<langSet xml:lang="en">
  <ntig>
    <termGrp>
      <term>state transition table</term>
    </termGrp>
  </ntig>
</langSet>
```

### 10.3.2 Synonyms

The following sample shows how synonyms can be represented in TBX. The following terminological data sample indicates that there is a synonym for the German term "Abtastglied":

- fr : échantillonneur
- en : sampling element
- de : Abtastglied; Abtaster

These terms are represented as follows:

```
<termEntry>
  <langSet xml:lang="en">
    <tig>
      <term>sampling element</term>
    </tig>
  </langSet>
  <langSet xml:lang="fr">
    <tig>
      <term>échantillonneur</term>
    </tig>
  </langSet>
  <langSet xml:lang="de">
    <tig>
      <term>Abtastglied</term>
    </tig>
    <tig>
      <term>Abtaster</term>
    </tig>
  </langSet>
</termEntry>
```

**NOTE** The use of the data-category */term type/* with a value of 'synonym' is not required to indicate synonyms, since two terms in the same language set are assumed to be synonyms unless otherwise indicated.

### 10.3.3 Abbreviations

The following samples show how abbreviations can be represented in TBX in two different methods. In the following terminological data sample the German term has an abbreviation:

- fr : élément à action proportionnelle et par intégration
- en : proportional plus integral element
- de : Proportionalglied plus Integrierglied; PI-Glied

The German langSet can be represented in TBX as:

```
<langSet xml:lang="de">
  <tig>
    <term>Proportionalglied plus Integrierglied</term>
  </tig>
  <tig>
    <term>PI-Glied</term>
    <termNote type="termType">abbreviation</termNote>
  </tig>
</langSet>
```

The German langSet can also be represented using "abbreviatedFormOf" as follows, when it is desirable to show the relationship between the abbreviated form and that full form explicitly:

```
<langSet xml:lang="de">
```

```

<tig>
  <term ID="n337">Proportionalglied plus Integrierglied</term>
</tig>
<tig>
  <term>PI-Glied</term>
  <termNote type="abbreviatedFormFor" target="n337">Proportionalglied plus
  Integrierglied</termNote>
</tig>
</langSet>

```

### 10.3.4 Transactions

The following is an example of how a terminological database workflow transaction can be recorded in TBX. The following type of construct can occur at any of the three levels of a terminological entry, to indicate workflow information for the entire concept, language section, or term section.

```

<transacGrp>
  <transac type="transactionType">modification</transac>
  <note>marketing department requested change from gizmo to thing-o</note>
  <date>1999-11-12</date>
  <transacNote type="responsibility">John Harris</transacNote>
</transacGrp>

```

### 10.3.5 Notes

The following is an example of how to record a note as well as the source of the note. The following type of construct can occur at any of the three levels of a terminological entry, to provide a note about the entire concept, language section, or term section.

```

<adminGrp>
  <admin type="annotatedNote">The Ashes is a Test cricket series, played between
  England and Australia.</admin>
  <adminNote type="noteSource">Wikipedia</adminNote>
</adminGrp>

```

If no source information for the note is required, use a simple <note> element, which can occur at many places in a terminological entry.

## 11 Referencing objects

### 11.1 General information about referencing

In TBX there are two methods of referencing objects. If the object is internal to the TBX document instance, it shall be embedded in the back matter of a TBX file. If the object is external to the TBX document instance, it shall be linked to by an <xref> element. In practice, most references are external.

Most terminological databases contain entries that need to reference objects that are external to the entry itself, such as a graphics file, an audio file, or a Web site. These objects may be accessible in the local environment (such as the computer file system) or from an external location such as a Web site on the Internet. TBX provides various ways to represent object references, which are shown in the following examples.

When referring to local files that are external to the TBX document instance using the *target* attribute on an <xref> element, it is necessary to specify the location of the files with a path that is relative to the directory containing the TBX document instance, rather than using an absolute path on a specific machine. For interchange purposes, it is also necessary to package the locally referenced folder and its contents with the TBX document instance. When using this method, it is recommended to include referenced objects such as graphic files and audio files in a designated directory.

When referencing a resource that is on the Internet, it is recommended to use a persistent identifier for the URI, to avoid links becoming broken when a URL for a Web site changes. Persistent identifiers called handles are described in The Handle System, available at [www.handle.net](http://www.handle.net).

Binary data embedded in the back matter of a TBX document instance shall be encoded in Base64.

The only type of referenced object defined in the TBX default XCS file is the binary object type. However, other types of referenced objects can be defined such as shown for bibliographic objects and people as in sections 11.6 and 11.7. Additional types may be defined in user-specific TBX TMLs. The XML elements chosen to represent those objects should be based on existing standards whenever possible.

NOTE The following examples focus on graphic files as objects. It should be noted that the `<xref>` element supports other *type* attribute values such as 'xAudio' for audio files.

## 11.2 Referencing a file that is embedded in the back matter of a TBX file

Referencing a file that is embedded in the back matter of a TBX file requires the use of a `<descrip>` element. In this example, the terminological entry contains a pointer to a graphic file that is embedded in the back matter in the form of binary data.

```
<body>
<termEntry>
...
  <descrip type="figure" target="fid-cricketbat">cricket bat</descrip>
...
</termEntry>
</body>
<back>
  <refObjectList type="binaryData">
    <refObject id="fid-cricketbat">
      <item type="codePage">base64</item>
      <item type="format">jpg</item>
      <item type="data">
/9j/4AAQSkZJRgABAQEAYABgAAD/2wBDAAAgGBgcGBQgHBwcJCQgKDBQNDAsLDBkSEw8UHRofHh0a...
</item>
    </refObject>
  </refObjectList>
</back>
```

## 11.3 Referencing a file from the back matter

This example shows how to record the location of the graphic in the back matter, rather than embedding the file in the back matter.

```
<body>
<termEntry>
...
  <descrip type="figure" target="fid-cricketbat">cricket bat</descrip>
...
</termEntry>
</body>
<back>
  <refObjectList type="binaryData">
    <refObject id="fid-cricketbat">
      <itemSet>
        <itemGrp>
          <item>bat.jpg</item>
          <xref target="sports/cricket/bat.jpg"/>
        </itemGrp>
      </itemSet>
    </refObject>
```

```

    </refObjectList>
</back>

```

## 11.4 Referencing a file directly in the entry

This example demonstrates how to use an `<xref>` element to specify the location of a file on the local file system directly in the entry, eliminating the need to include information in the back matter.

```

<termEntry>
...
  <xref type="xGraphic" target="sports/cricket/bat.jpg">cricket bat</xref>
...
</termEntry>

```

## 11.5 Referencing an external source

The following example adds a reference to a Web site from where the graphic was originally obtained.

```

<termEntry>
...
  <xref type="xGraphic" target="sports/cricket/bat.jpg">cricket bat</xref>
  <xref type="xSource" target="http://www.bestsportsgraphics.com">
    Best Sports Graphics</xref>
...
</termEntry>

```

This type of reference can also be used for definitions, context sentences, notes, and so forth, as the following example demonstrates. This example shows the use of a `<descripGrp>` element to group together a definition and its external source.

```

<termEntry>
<langSet xml:id="en-US">
  <tig>
    <term>over</term>
    <descripGrp>
      <descrip type="definition">Six balls bowled in sequence by the same
        bowler.</descrip>
      <xref type="xSource" target="http://www.allaboutcricket.com">
        All About Cricket</xref>
    </descripGrp>
  </tig>
</langSet>
</termEntry>

```

## 11.6 Referencing and documenting a bibliographic source

The following example shows how to reference, in a terminological entry, a bibliographic source that is documented in the back matter of a TBX document instance. Data-categories for bibliographic references (values of the *type* attribute on the `<item>` elements) are adopted from the Dublin Core Metadata Element Set. The example given below only includes a subset of the available data-categories for demonstration purposes. The XML structure for bibliographic references is adopted from the one already specified in the default XCS file for binaryData.

```

<body>
<termEntry>

```

```

<langSet xml:lang="en-US">
  <tig>
    <term>over</term>
    <descripGrp>
      <descrip type="definition">Six balls bowled in sequence by the same
      bowler.</descrip>
      <admin type="sourceIdentifier" target="bid-cricket">The Complete Guide
      to Cricket, p. 17</admin>
    </descripGrp>
  </tig>
</langSet>
</termEntry>
</body>
<back>
  <refObjectList type="bibl">
    <refObject id="bid-cricket">
      <item type="creator">Farrel, Bart</item>
      <item type="title">The Complete Guide to Cricket</item>
      <item type="type">book</item>
      <item type="identifier">978-3-16-148410-0</item>
      <item type="publisher">Sports Publishers Inc.</item>
      <item type="date">2000</item>
    </refObject>
  </refObjectList>
</back>

```

## 11.7 Referencing and documenting information about a responsible person or organization

The following example shows how to reference information about a person and an organization who are responsible for part of a terminological entry, which is documented in the back matter of a TBX document instance. Data-categories for personal or organizational references (values of the *type* attribute on the `<item>` elements) are adopted from the vCard Standard. The example given below only includes a subset of the available data-categories for demonstration purposes. The XML structure for these references in the back matter is adopted from the one already specified in the default XCS file for 'binaryData'.

```

<body>
<termEntry>
<langSet xml:lang="en-US">
  <tig>
    <term>rabbit</term>
    <descripGrp>
      <descrip type="definition">In the game of cricket, a very poor batsman.
      </descrip>
      <transacGrp>
        <transac type="transactionType">importation</transac>
        <transacNote type="responsibility" target="rpid-harris">John
        Harris</transacNote>
      </transacGrp>
      <transacGrp>
        <transac type="transactionType">origination</transac>
        <transacNote type="responsibility" target="roid-CCA">Canadian
        Cricket Association</transacNote>
      </transacGrp>
    </descripGrp>
  </tig>
</langSet>
</termEntry>
</body>
<back>
  <refObjectList type="respPerson">

```

```

    <refObject id="rpid-harris">
      <item type="fn">John Harris</item>
      <item type="title">Senior terminologist</item>
      <item type="role">Approver</item>
      <item type="email">jharris@myserver.com</item>
      <item type="uid">xyz123</item>
    </refObject>
  </refObjectList>
  <refObjectList type="respOrg">
    <refObject id="roid-CCA">
      <item type="org">Canadian Cricket Association</item>
      <item type="tel">123-456-7899</item>
      <item type="email">cca@anotherserver.com</item>
      <item type="adr">75 1st Street, Ottawa, Ontario, Canada</item>
    </refObject>
  </refObjectList>
</back>

```

## 11.8 Referencing an external concept system, classification system, or thesaurus

An external concept system, classification system, or thesaurus can sometimes be the source of information about concepts. In this case it is recommended to point to the external resource by using an `<xref>` element with a unique URI, rather than documenting information about that source in the back matter of the TBX document instance.

## 11.9 Referencing a TBX entry from within a corpus

The Localization Industry Standards Association has published TermLink, a specification for referencing terminological database entries, which is available at [www.lisa.org/Term-Link.109.0.html](http://www.lisa.org/Term-Link.109.0.html).

## 12 Creating customized TBX TMLs

### 12.1 General information about TMLs

Few terminology collections or applications use the same set of data-categories. TBX is a flexible format because it allows user groups to select their own data-categories. By doing so, they can create their own TML adapted to their requirements. The data-categories and their constraints are represented in an XCS file.

If the data-categories selected by a user group are already found in the default TBX XCS file, the TML is a strict subset of the TBX-default TML. If a user group selects all the data-categories that are in the default TBX XCS file, plus additional ones, the TML is a strict superset of the TBX-default TML. A user group may also select some of the data categories from the default TBX TML plus some additional ones, to create a mixed set. To facilitate interoperability, any data-categories used that are not part of the default TBX XCS file shall be described in comments in the header of the user-specific XCS file. In addition, if any of the additional data-categories are not already available in ISO 12620:1999, there should be a statement in the header of the XCS file indicating how and when they will be submitted to ISO TC37.

Whenever possible, data-categories shall be selected from ISO 12620. TBX TMLs are interoperable and interchangeable whether they are a superset or a subset of the TBX-default TML, with possibly a certain level of negotiation required to prevent or minimize data loss. TMLs that are a subset of the TBX-default TML will obviously be easier to interchange and require less negotiation than superset TMLs.

The following figure shows how the flexibility of TBX is based on the classic form-content distinction. Each combination of the core structure DTD/schema, which defines the form, and a particular XCS file, which defines the allowed content, results in a customized variant (TML) of TBX. Each TBX TML complies with TMF (ISO 16642) because it has the same core structure as TBX. It may differ from other TBX TMLs with respect to the data-categories and constraints on those data-categories.

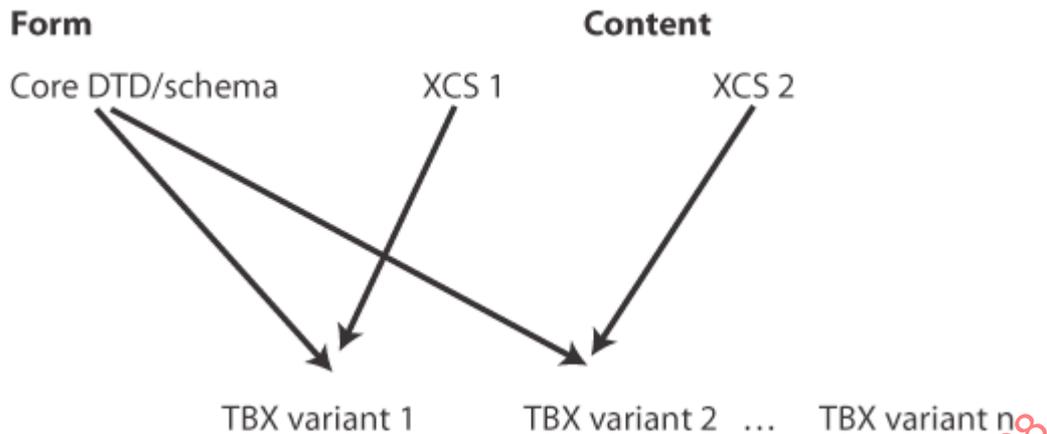


Figure 4. Creating TBX TMLs

In addition to the TBX-default TML, one TBX variant (TML) is referred to in this International Standard: [TBX Basic](#).

To demonstrate how to create a user-specific TBX TML, this section will describe a fictitious TBX TML. For illustrative purposes, this TML will allow minimal terminological information provided to a fictitious supplier of translation along with the source text to be translated, in the fields of manufacturing and finance.

This TML will allow only two types of terms, full forms and abbreviated forms. This is achieved by specifying a picklist of permissible values (sometimes called a value domain) for the data-category */term type/*, which is an instantiation of the meta data-category `<termNote>`. The following information is placed in the XCS module concerning term type:

- meta data-category: `<termNote>`
- data-category: */term type/*
- picklist: fullForm, abbreviatedForm

This specification is a strict subset of the specification for *termType* in the default XCS file. The only difference is that the default XCS file allows more options in the picklist. Clearly, any instance document that conforms to this TML also conforms to TBX.

This TML will allow only two types of descriptive information: a subject field and a definition. The subject fields allowed in this subset are manufacturing and finance, and subject field specifications are allowed only at the terminological entry level:

- meta data-category: `<descrip>`
- data-category: */subject field/*
- picklist: manufacturing, finance
- levels: termEntry

The default XCS file allows any plainText value for a subject field, so a subset XCS file can specify a picklist. Obviously, there can be no picklist of possible definitions, so the specification for the definition contains the same type of text found in general notes (*noteText*) and is allowed at two levels, entry and language.

This is done by placing the following information in the subset XCS file:

- meta data-category: `<descrip>`
- data-category: */definition/*
- content: noteText
- levels: termEntry, langSet

## 12.2 Example of an XCS file for a user-defined TBX TML

An XCS file for the TML described in the previous section would look like this:

```

<xml version="1.0"?>
<!DOCTYPE TBXXCS SYSTEM "tbxxcsdtd.dtd">
<TBXXCS id="XCS-supplier" version="1.0" lang="en">
<header>

```

```

<title>subset XCS file for suppliers</title>
</header>
<datCatSet>
  <termNoteSpec id="termType" datcatId="ISO12620A-0201">
    <contents datatype="picklist">fullForm
    abbreviatedForm</contents>
  </termNoteSpec>
  <descripSpec id="subjectField" datcatId="ISO12620A-04">
    <contents datatype="picklist">manufacturing
    finance</contents>
    <levels>termEntry</levels>
  </descripSpec>
  <descripSpec id="definition" datcatId="ISO12620A-0501">
    <contents/>
    <levels>termEntry langSet</levels>
  </descripSpec>
</datCatSet>
</TBXXCS>

```

NOTE For simplicity purposes, the elements `<languages>` and `<refObjectDefSet>` have been omitted in the above sample.

It should be noted that the machine-processable XCS module corresponds to the information listed for the three data-categories presented in the above examples. The data-category identifiers (such as ISO12620A-02010) are found in the default XCS file.

Clearly, specifying only three data-categories (term type, subject field, and definition) as instances of meta data-categories defines a very limited subset of TBX; nevertheless, this limited data-category module can be logically combined with the core-structure module of TBX to allow such TBX-compliant instances as the one in [10.1 Example of a typical TBX file](#). Elements that are not meta data-categories, i.e., basic TBX data-categories such as `<term>` and `<note>`, are not specified in the XCS module because they are part of the core structure.

### 12.3 Creating customized picklist display names

The permissible picklist values for data-categories are specified in the [Annex C \(Normative\) Default XCS file](#), as the content of the corresponding data-category element, for example, `administrativeStatus`. The representation style of these picklist tokens is camelcase, such as `standardizedTerm`. An end-user can document customized display names for these tokens, by using the following elements in the XCS header:

`<datCatDoc>`

The root element of the section in the XCS header that contains picklist value mappings. It contains zero or more `<datCatMap>` elements.

`<datCatMap>`

An element that encloses one picklist value mapping. It contains one `<datCatToken>`, one or more `<datCatDisplay>`, and optionally one or more `<langCode>` and `<datCatNote>`.

`<datCatToken>`

The name of the picklist value in the XCS file, i.e., one of the values in the `<contents>` element.

`<datCatDisplay>`

The user-specific display name for the picklist value.

`<langCode>`

(Optional) A language to which the data-category applies. Can be repeated to indicate multiple languages.

`<datCatNote>`

(Optional) Any note about the data-category.

These elements can be used to create a mapping between user-specified display names and the data-category

names. This mechanism is useful for applications where, for example, the data-category needs to be presented to the user in a language other than English, or where an application prefers a presentation style other than camel case.

For example, the following markup in the XCS header specifies alternate names for the gender picklist values in French for a terminological database that has a French user interface:

```
<header>
  <title>XCS file for a French termbase</title>
  <datCatDoc>
    <datCatMap>
      <datCatToken>masculine</datCatToken>
      <datCatDisplay>masculin</datCatDisplay>
      <langCode>fr-FR</langCode>
      <langCode>fr-CA</langCode>
    </datCatMap>
    <datCatMap>
      <datCatToken>feminine</datCatToken>
      <datCatDisplay>féminin</datCatDisplay>
      <langCode>fr-FR</langCode>
      <langCode>fr-CA</langCode>
    </datCatMap>
    <datCatMap>
      <datCatToken>neuter</datCatToken>
      <datCatDisplay>neutre</datCatDisplay>
      <langCode>fr-FR</langCode>
      <langCode>fr-CA</langCode>
      <datCatNote xml:lang="fr">Ne s'applique qu'aux langues comme
        l'allemand.</datCatNote>
    </datCatMap>
    <datCatMap>
      <datCatToken>otherGender</datCatToken>
      <datCatDisplay>autre genre</datCatDisplay>
      <langCode>fr-FR</langCode>
      <langCode>fr-CA</langCode>
    </datCatMap>
  </datCatDoc>
</header>
```

STANDARDSISO.COM : Click to view the full PDF of ISO 30042:2008

## Annex A (Normative) DTD for the core structure module

This annex contains a formal representation of the core structure as a DTD. When reformatted as a separate file, it should be named *TBXcoreStructV02.dtd* (TBX Core DTD Version 02).

The XML entities (such as `noteText`) listed in the DTD allow mnemonic names to be given to text strings, especially text that is used in several places. The elements of TBX are divided into three groups: (a) the low-level elements used to mark up text, such as markup inside definitions and contextual examples, (b) elements needed to constitute a terminological entry (<[termEntry](#)>), and (c) high-level elements and other elements not used in a terminological entry, such as header elements.

```

<!-- TBXcoreStructV02.dtd -->
<!-- note: see XCS for values of type on meta data-categories and for values of
xml:lang -->

<!-- =====
SOME USEFUL ENTITIES THAT ARE REFERENCED BELOW
===== -->
<!ENTITY % basicText '(#PCDATA | hi)*'>
<!ENTITY % noteText '(#PCDATA | hi | foreign | bpt | ept | ph)*'>
<!ENTITY % auxInfo '(descrip | descripGrp | admin | adminGrp | transacGrp | note | ref
| xref)*' >
<!ENTITY % noteLinkInfo '(admin | adminGrp | transacGrp | note | ref | xref)*' >
<!-- Entities that define common sets of attributes -->
<!ENTITY % impIDLang '
  id      ID #IMPLIED
  xml:lang CDATA #IMPLIED
'>
<!ENTITY % impIDType '
  id      ID #IMPLIED
  type    CDATA #IMPLIED
'>
<!ENTITY % impIDLangTypTgtDtyp '
  id      ID #IMPLIED
  xml:lang CDATA #IMPLIED
  type    CDATA #REQUIRED
  target  IDREF #IMPLIED
  datatype CDATA #IMPLIED
'>

<!-- =====
ELEMENTS USED FOR TEXT MARKUP
===== -->
<!ELEMENT hi      (#PCDATA) >
<!ATTLIST hi
  type    CDATA #IMPLIED
  target  IDREF #IMPLIED
  xml:lang CDATA #IMPLIED
>

<!ELEMENT foreign %noteText; >
<!ATTLIST foreign
  id      ID #IMPLIED
  xml:lang CDATA #IMPLIED
>

<!ELEMENT bpt (#PCDATA) >
<!ATTLIST bpt
  i      CDATA #IMPLIED
  type   CDATA #IMPLIED
>

```

```

<!ELEMENT ept (#PCDATA) >
<!ATTLIST ept
  i      CDATA #IMPLIED
>

<!ELEMENT ph (#PCDATA) >
<!ATTLIST ph
  type   CDATA #IMPLIED
>

<!-- =====
ELEMENTS NEEDED FOR TERMINOLOGICAL ENTRIES (IN ALPHABETICAL ORDER)
===== -->

<!ELEMENT admin %noteText; >
<!ATTLIST admin
  %impIDLangTypTgtDtyp;
>

<!ELEMENT adminGrp (admin, (adminNote|note|ref|xref)*) >
<!ATTLIST adminGrp
  id      ID #IMPLIED
>

<!ELEMENT adminNote (#PCDATA) >
<!ATTLIST adminNote
  %impIDLangTypTgtDtyp;
>

<!ELEMENT date (#PCDATA) >
<!ATTLIST date
  id      ID #IMPLIED
>

<!ELEMENT descrip %noteText; >
<!ATTLIST descrip
  %impIDLangTypTgtDtyp;
>

<!ELEMENT descripGrp (descrip, (descripNote|admin|adminGrp|transacGrp|note|ref|xref)*) >
<!ATTLIST descripGrp
  id      ID #IMPLIED
>

<!ELEMENT descripNote (#PCDATA) >
<!ATTLIST descripNote
  %impIDLangTypTgtDtyp;
>

<!ELEMENT langSet ((%auxInfo;), (tig | ntig)+) >
<!ATTLIST langSet
  id      ID #IMPLIED
  xml:lang CDATA #REQUIRED
>

<!ELEMENT note %noteText; >
<!ATTLIST note %impIDLang;
>

<!ELEMENT ntig (termGrp, %auxInfo;) >
<!ATTLIST ntig
  id      ID #IMPLIED

```

```

>
<!ELEMENT ref (#PCDATA) >
<!ATTLIST ref
  %impIDLangTypTgtDtyp;
>

<!ELEMENT term %basicText; >
<!ATTLIST term
  id      ID #IMPLIED
>

<!ELEMENT termComp (#PCDATA) >
<!ATTLIST termComp
  %impIDLang;
>

<!ELEMENT termCompGrp (termComp, (termNote|termNoteGrp)*, %noteLinkInfo;) >
<!ATTLIST termCompGrp
  id      ID #IMPLIED
>

<!ELEMENT termCompList ((%auxInfo;), (termComp | termCompGrp)+) >
<!ATTLIST termCompList
  id      ID #IMPLIED
  type    CDATA #REQUIRED
>

<!ELEMENT termEntry ((%auxInfo;), (langSet+)) >
<!ATTLIST termEntry
  id      ID #IMPLIED
>

<!ELEMENT termGrp (term, (termNote|termNoteGrp)*, (termCompList)* ) >
<!ATTLIST termGrp
  id      ID #IMPLIED
>

<!ELEMENT termNote %noteText; >
<!ATTLIST termNote
  %impIDLangTypTgtDtyp;
>

<!ELEMENT termNoteGrp (termNote, %noteLinkInfo;) >
<!ATTLIST termNoteGrp
  id      ID #IMPLIED
>

<!ELEMENT tig (term, (termNote)*, %auxInfo;) >
<!ATTLIST tig
  id      ID #IMPLIED
>

<!ELEMENT transac (#PCDATA) >
<!ATTLIST transac
  %impIDLangTypTgtDtyp;
>

<!ELEMENT transacGrp (transac, (transacNote|date|note|ref|xref)* ) >
<!ATTLIST transacGrp
  id      ID #IMPLIED
>

```

## ISO 30042:2008(E)

```
<!ELEMENT transacNote (#PCDATA) >
<!ATTLIST transacNote
  %impIDLangTypTgtDtyp;
>

<!ELEMENT xref (#PCDATA) >
<!ATTLIST xref
  %impIDType;
  target CDATA #REQUIRED
>

<!--
=====
OTHER ELEMENTS (in hierarchical order)
=====
-->
<!ELEMENT martif (martifHeader, text) > <!-- *** starting element *** -->
<!ATTLIST martif
  type CDATA #REQUIRED
  xml:lang CDATA #REQUIRED
>

<!ELEMENT martifHeader (fileDesc, encodingDesc?, revisionDesc?) >
<!ATTLIST martifHeader
  id ID #IMPLIED
>

<!ELEMENT p %noteText; > <!-- p is used in several header elements -->
<!ATTLIST p
  id ID #IMPLIED
  type (DCSName|XCSURI|XCSCContent) #IMPLIED
  xml:lang CDATA #IMPLIED
>

<!ELEMENT fileDesc (titleStmnt?, publicationStmnt?, sourceDesc+) >
<!ATTLIST fileDesc
  id ID #IMPLIED
>

<!ELEMENT titleStmnt (title, note*) >
<!ATTLIST titleStmnt
  %impIDLang;
>

<!ELEMENT title (#PCDATA) >
<!ATTLIST title
  %impIDLang;
>

<!ELEMENT publicationStmnt (p+) >
<!ATTLIST publicationStmnt
  id ID #IMPLIED
>

<!ELEMENT sourceDesc (p+) >
<!ATTLIST sourceDesc
  %impIDLang;
>

<!ELEMENT encodingDesc (p+) >
<!ATTLIST encodingDesc
  id ID #IMPLIED
>
```

```

<!ELEMENT revisionDesc (change+) >
<!ATTLIST revisionDesc
  %impIDLang;
>

<!ELEMENT change (p+) >
<!ATTLIST change
  %impIDLang; >

<!ELEMENT text (body, back?) >
<!ATTLIST text
  id ID #IMPLIED
>

<!ELEMENT body (termEntry+) >
<!ATTLIST body
  id ID #IMPLIED
>

<!ELEMENT back ((refObjectList)*) >
<!ATTLIST back
  id ID #IMPLIED
>

<!ELEMENT refObjectList (refObject+) >
<!ATTLIST refObjectList
  id ID #IMPLIED
  type CDATA #REQUIRED
>

<!ELEMENT refObject ((itemSet | itemGrp | item)+) >
<!ATTLIST refObject
  id ID #IMPLIED
>

<!ELEMENT item %noteText; >
<!ATTLIST item
  %impIDType;
>

<!ELEMENT itemGrp (item, %noteLinkInfo;)>
<!ATTLIST itemGrp
  id ID #IMPLIED
>

<!ELEMENT itemSet ((item | itemGrp)+)>
<!ATTLIST itemSet
  %impIDType;
>

<!-- end -->

```

## Annex B (Normative) DTD for the data-category constraints (XCS file)

The following is the DTD that can be used to validate the default XCS file, or any user-specific XCS file representing a subset of the default XCS file. The datatype plainText is used in the default XCS file but is not defined in the following DTD. The value plainText refers to PCDATA.

```

<!-- start classes -->
<!ENTITY % specAtt.attributes '
datcatId CDATA #REQUIRED
name CDATA #REQUIRED'>
<!-- stop classes -->

<!-- start elements -->
<!ELEMENT TBXXCS ((header, languages, datCatSet), (refObjectDefSet?))>
<!ATTLIST TBXXCS
lang CDATA #REQUIRED
name CDATA #REQUIRED
version CDATA #REQUIRED >

<!ELEMENT adminNoteSpec (contents)>
<!ATTLIST adminNoteSpec
%specAtt.attributes; >

<!ELEMENT adminSpec (contents)>
<!ATTLIST adminSpec
%specAtt.attributes; >

<!ELEMENT contents (#PCDATA)>
<!ATTLIST contents
datatype CDATA #IMPLIED
forTermComp CDATA #IMPLIED
targetType CDATA #IMPLIED >

<!ELEMENT datCatDisplay (#PCDATA)>
<!ATTLIST datCatDisplay
xml:lang CDATA #IMPLIED >

<!ELEMENT datCatDoc (datCatMap)*>

<!ELEMENT datCatMap (datCatToken, (datCatDisplay)+, (langCode)*, (datCatNote)?)>

<!ELEMENT datCatNote (#PCDATA)>
<!ATTLIST datCatNote
xml:lang CDATA #IMPLIED >

<!ELEMENT datCatSet (adminNoteSpec |
adminSpec |
descripNoteSpec |
descripSpec |
hiSpec |
refSpec |
termCompListSpec |
termNoteSpec |
transacNoteSpec |
transacSpec |
xrefSpec)+>

<!ELEMENT datCatToken (#PCDATA)>

<!ELEMENT descripNoteSpec (contents)>
<!ATTLIST descripNoteSpec

```

```

%specAtt.attributes; >

<!ELEMENT descripSpec (contents,levels)>
<!ATTLIST descripSpec
  %specAtt.attributes; >

<!ELEMENT header (title,(datCatDoc)?)>

<!ELEMENT hiSpec (contents)>
<!ATTLIST hiSpec
  %specAtt.attributes; >

<!ELEMENT itemSpec (#PCDATA)>
<!ATTLIST itemSpec
  type CDATA #REQUIRED >

<!ELEMENT itemSpecSet (itemSpec)+>
<!ATTLIST itemSpecSet
  type CDATA #REQUIRED >

<!ELEMENT langCode (#PCDATA)>

<!ELEMENT langInfo (langCode,langName)>

<!ELEMENT langName (#PCDATA)>

<!ELEMENT languages (langInfo)+>

<!ELEMENT levels (#PCDATA)>

<!ELEMENT refObjectDef (refObjectType,itemSpecSet)>

<!ELEMENT refObjectDefSet (refObjectDef)+>

<!ELEMENT refObjectType (#PCDATA)>

<!ELEMENT refSpec (contents)>
<!ATTLIST refSpec
  %specAtt.attributes; >

<!ELEMENT termCompListSpec (contents)>
<!ATTLIST termCompListSpec
  %specAtt.attributes; >

<!ELEMENT termNoteSpec (contents)>
<!ATTLIST termNoteSpec
  %specAtt.attributes; >

<!ELEMENT title (#PCDATA)>

<!ELEMENT transacNoteSpec (contents)>
<!ATTLIST transacNoteSpec
  %specAtt.attributes; >

<!ELEMENT transacSpec (contents)>
<!ATTLIST transacSpec
  %specAtt.attributes; >

<!ELEMENT xrefSpec (contents)>
<!ATTLIST xrefSpec
  %specAtt.attributes; >
<!-- end elements -->

```

## Annex C (Normative) Default XCS file

### C.1 Introduction

This annex contains a machine processable definition of the default selection of data-categories for TBX and the constraints on those data-categories, in the form of an XCS file. The XCS file applies the data-category constraints to the core structure. Specific user-group implementations of TBX may use a different XCS file, as described in [12 Creating customized TBX TMLs](#). A description of the components of the XCS file is provided in [Annex E](#). A DTD that can be used to validate any XCS file for TBX TMLs is provided in [Annex B](#).

The conditions under which this annex is normative is explained in clause 7.1.

An open source application that checks whether a given TBX document instance adheres to the XCS specified in its header is available from the Localization Industry Standards Association. For more information, see [Annex I.2 External references](#).

### C.2 XCS file for the default data-categories and constraints

NOTE The datatype plainText used in the default XCS file refers to PCDATA.

```
<?xml version="1.0" encoding="UTF-8"?> <!DOCTYPE TBXXCS SYSTEM "tbxxcsdtd.dtd">
<TBXXCS name="master" version="0.4" lang="en">
  <header>
    <title>TBX default XCS (extensible constraint specification) file Version 02.
    (Filename: TBXXCSV02.XCS)</title>
    <!--location where, if the default XCS file is modified to include additional data-
    categories, those additions are also listed and described here for convenience
    purposes-->
  </header>
  <languages>
    <langInfo>
      <langCode>en</langCode>
      <langName>English</langName>
    </langInfo>
    <langInfo>
      <langCode>de</langCode>
      <langName>German</langName>
    </langInfo>
  </languages>
  <datCatSet>
    <!--data-category constraint specifications are arranged in alphabetical order by the
    value of the name attribute-->
    <termNoteSpec name="abbreviatedFormFor" datcatId="ISO12620A-02013002">
      <contents datatype="basicText" targetType="term"/>
    </termNoteSpec>
    <termNoteSpec name="administrativeStatus" datcatId="ISO12620A-020903">
      <contents datatype="picklist">
        standardizedTerm-admn-sts preferredTerm-admn-sts
        admittedTerm-admn-sts deprecatedTerm-admn-sts supersededTerm-admn-sts legalTerm-
        admn-sts regulatedTerm-admn-sts </contents>
      </termNoteSpec>
    <termNoteSpec name="animacy" datcatId="ISO12620A-020204">
      <contents datatype="picklist" forTermComp="yes">animate inanimate
      otherAnimacy</contents>
    </termNoteSpec>
    <adminSpec name="annotatedNote" datcatId="">
      <contents/>
    </adminSpec>
    <descripSpec name="antonymConcept" datcatId="ISO12620A-10180602">
```

```

    <contents datatype="basicText" targetType="entry"/>
  </levels>termEntry</levels>
</descripSpec>
<termNoteSpec name="antonymTerm" datcatId="ISO12620A-10180601">
  <contents datatype="basicText" targetType="term"/>
</termNoteSpec>
<adminSpec name="applicationSubset" datcatId="ISO12620A-100306">
  <contents datatype="plainText"/>
</adminSpec>
<descripSpec name="associatedConcept" datcatId="ISO12620A-070210">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="audio" datcatId="ISO12620A-050502">
  <contents datatype="plainText" targetType="binaryData"/>
  <levels>langSet termEntry term</levels>
</descripSpec>
<hiSpec name="bold" datcatId="">
  <contents/>
</hiSpec>
<descripSpec name="broaderConceptGeneric" datcatId="ISO12620A-070201">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="broaderConceptPartitive" datcatId="ISO12620A-070202">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<adminSpec name="businessUnitSubset" datcatId="ISO12620A-100308">
  <contents datatype="plainText"/>
</adminSpec>
<descripSpec name="characteristic" datcatId="ISO12620A-0508">
  <contents datatype="plainText"/>
  <levels>term</levels>
</descripSpec>
<descripSpec name="classificationCode" datcatId="ISO12620A-0402">
  <contents datatype="plainText" targetType="bibl"/>
  <levels>langSet termEntry term</levels>
</descripSpec>
<adminSpec name="conceptOrigin" datcatId="ISO12620A-0509">
  <contents datatype="plainText"/>
</adminSpec>
<descripSpec name="conceptPosition" datcatId="ISO12620A-0702">
  <contents datatype="plainText" targetType="conceptSysDescrip"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="context" datcatId="ISO12620A-0503">
  <contents/>
  <levels>term</levels>
</descripSpec>
<descripNoteSpec name="contextType" datcatId="ISO12620A-0803">
  <contents datatype="picklist">definingContext explanatoryContext associativeContext
linguisticContext metalinguisticContext translatedContext   </contents>
</descripNoteSpec>
<descripSpec name="coordinateConceptGeneric" datcatId="ISO12620A-07020401">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="coordinateConceptPartitive" datcatId="ISO12620A-07020402">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<xrefSpec name="corpusTrace" datcatId="ISO12620A-101808">

```

```

    <contents targetType="external"/>
  </xrefSpec>
  <refSpec name="crossReference" datcatId="ISO12620A-1018">
    <contents targetType="element"/>
  </refSpec>
  <adminSpec name="customerSubset" datcatId="ISO12620A-100301">
    <contents datatype="plainText"/>
  </adminSpec>
  <adminSpec name="databaseType" datcatId="ISO12620A-10220301">
    <contents datatype="plainText"/>
  </adminSpec>
  <descripSpec name="definition" datcatId="ISO12620A-0501">
    <contents/>
    <levels>langSet termEntry term</levels>
  </descripSpec>
  <descripNoteSpec name="definitionType" datcatId="ISO12620A-0802">
    <contents datatype="picklist">intensionalDefinition extensionalDefinition
    partitiveDefinition translatedDefinition</contents>
  </descripNoteSpec>
  <descripNoteSpec name="descriptionType" datcatId="ISO12620A-0801">
    <contents datatype="picklist" targetType="element"/>
  </descripNoteSpec>
  <termNoteSpec name="directionality" datcatId="">
    <contents datatype="picklist" targetType="term">monodirectional bidirectional
    incommensurate undetermined</contents>
  </termNoteSpec>
  <adminSpec name="domainExpert" datcatId="ISO12620A-1025">
    <contents datatype="plainText" targetType="bibl"/>
  </adminSpec>
  <adminSpec name="elementWorkingStatus" datcatId="ISO12620A-1011">
    <contents datatype="picklist">starterElement workingElement consolidatedElement
    archiveElement importedElement exportedElement</contents>
  </adminSpec>
  <hiSpec name="entailedTerm" datcatId="ISO12620A-100601">
    <contents/>
  </hiSpec>
  <adminSpec name="entrySource" datcatId="ISO12620A-1013">
    <contents datatype="plainText"/>
  </adminSpec>
  <adminSpec name="environmentSubset" datcatId="ISO12620A-100307">
    <contents datatype="plainText"/>
  </adminSpec>
  <termNoteSpec name="etymology" datcatId="ISO12620A-020402">
    <contents forTermComp="yes"/>
  </termNoteSpec>
  <descripSpec name="example" datcatId="ISO12620A-0504">
    <contents/>
    <levels>langSet termEntry term</levels>
  </descripSpec>
  <descripSpec name="explanation" datcatId="ISO12620A-0502">
    <contents/>
    <levels>langSet termEntry term</levels>
  </descripSpec>
  <xrefSpec name="externalCrossReference" datcatId="ISO12620A-101807">
    <contents targetType="external"/>
  </xrefSpec>
  <termNoteSpec name="falseFriend" datcatId="ISO12620A-0302">
    <contents datatype="basicText" targetType="term"/>
  </termNoteSpec>
  <descripSpec name="figure" datcatId="ISO12620A-050501">
    <contents datatype="plainText" targetType="binaryData"/>
    <levels>langSet termEntry term</levels>
  </descripSpec>

```

```

<termNoteSpec name="frequency" datcatId="ISO12620A-020304">
  <contents datatype="picklist">commonlyUsed infrequentlyUsed rarelyUsed</contents>
</termNoteSpec>
<termNoteSpec name="geographicalUsage" datcatId="ISO12620A-020302">
  <contents datatype="plainText"/>
</termNoteSpec>
<termNoteSpec name="grammaticalGender" datcatId="ISO12620A-020202">
  <contents datatype="picklist" forTermComp="yes">masculine feminine neuter
otherGender</contents>
</termNoteSpec>
<termNoteSpec name="grammaticalNumber" datcatId="ISO12620A-020203">
  <contents datatype="picklist" forTermComp="yes">singular plural dual mass
otherNumber</contents>
</termNoteSpec>
<termNoteSpec name="grammaticalValency" datcatId="ISO12620A-020207">
  <contents datatype="plainText"/>
</termNoteSpec>
<termNoteSpec name="homograph" datcatId="ISO12620A-101805">
  <contents datatype="basicText" targetType="term"/>
</termNoteSpec>
<hiSpec name="hotkey" datcatId="ISO12620A-100604">
  <contents/>
</hiSpec>
<termCompListSpec name="hyphenation" datcatId="ISO12620A-0207">
  <contents forTermComp="yes"/>
</termCompListSpec>
<adminSpec name="indexHeading" datcatId="ISO12620A-0905">
  <contents datatype="plainText"/>
</adminSpec>
<hiSpec name="italics" datcatId="">
  <contents/>
</hiSpec>
<adminSpec name="keyword" datcatId="ISO12620A-0904">
  <contents datatype="plainText"/>
</adminSpec>
<termNoteSpec name="language-planningQualifier" datcatId="ISO12620A-020902">
  <contents datatype="picklist">recommendedTerm nonstandardizedTerm proposedTerm
newTerm</contents>
</termNoteSpec>
<termCompListSpec name="lemma" datcatId="ISO12620A-020803">
  <contents forTermComp="yes"/>
</termCompListSpec>
<termNoteSpec name="lionHotkey" datcatId="">
  <contents datatype="plainText"/>
</termNoteSpec>
<hiSpec name="math" datcatId="">
  <contents/>
</hiSpec>
<termCompListSpec name="morphologicalElement" datcatId="ISO12620A-020801">
  <contents forTermComp="yes"/>
</termCompListSpec>
<termNoteSpec name="normativeAuthorization" datcatId="ISO12620A-020901">
  <contents datatype="picklist">standardizedTerm preferredTerm admittedTerm
deprecatedTerm supersededTerm legalTerm regulatedTerm</contents>
</termNoteSpec>
<adminNoteSpec name="noteSource" datcatId="">
  <contents/>
</adminNoteSpec>
<adminSpec name="originatingDatabase" datcatId="ISO12620A-102203">
  <contents datatype="plainText"/>
</adminSpec>
<adminSpec name="originatingInstitution" datcatId="ISO12620A-102202">
  <contents datatype="plainText"/>

```

```

</adminSpec>
<adminSpec name="originatingPerson" datcatId="ISO12620A-102201">
  <contents datatype="plainText"/>
</adminSpec>
<descripSpec name="otherBinaryData" datcatId="ISO12620A-050505">
  <contents datatype="plainText" targetType="binaryData"/>
  <levels>langSet termEntry term</levels>
</descripSpec>
<termNoteSpec name="partOfSpeech" datcatId="ISO12620A-020201">
  <contents datatype="plainText" forTermComp="yes"/>
</termNoteSpec>
<termNoteSpec name="processStatus" datcatId="ISO12620A-020904">
  <contents datatype="picklist">unprocessed provisionallyProcessed
finalized</contents>
</termNoteSpec>
<adminSpec name="productSubset" datcatId="ISO12620A-100305">
  <contents datatype="plainText"/>
</adminSpec>
<adminSpec name="projectSubset" datcatId="ISO12620A-100303">
  <contents datatype="plainText"/>
</adminSpec>
<termNoteSpec name="pronunciation" datcatId="ISO12620A-020505">
  <contents datatype="basicText" forTermComp="yes"/>
</termNoteSpec>
<termNoteSpec name="proprietaryRestriction" datcatId="ISO12620A-020307">
  <contents datatype="picklist">trademark serviceMark tradeName</contents>
</termNoteSpec>
<descripSpec name="quantity" datcatId="ISO12620A-050701">
  <contents datatype="plainText"/>
  <levels>term</levels>
</descripSpec>
<descripSpec name="range" datcatId="ISO12620A-050707">
  <contents datatype="plainText"/>
  <levels>term</levels>
</descripSpec>
<termNoteSpec name="register" datcatId="ISO12620A-020303">
  <contents datatype="picklist">colloquialRegister neutralRegister technicalRegister
in-houseRegister bench-levelRegister slangRegister vulgarRegister</contents>
</termNoteSpec>
<descripSpec name="relatedConcept" datcatId="ISO12620A-070205">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="relatedConceptBroader" datcatId="ISO12620A-07020501">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="relatedConceptNarrower" datcatId="ISO12620A-07020502">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="reliabilityCode" datcatId="ISO12620A-0304">
  <contents datatype="picklist">1 2 3 4 5 6 7 8 9 10</contents>
  <levels>langSet termEntry term</levels>
</descripSpec>
<transacNoteSpec name="responsibility" datcatId="ISO12620A-100202">
  <contents targetType="respPerson"/>
</transacNoteSpec>
<descripSpec name="sampleSentence" datcatId="ISO12620A-050302">
  <contents/>
  <levels>term</levels>
</descripSpec>
<adminSpec name="searchTerm" datcatId="ISO12620A-100603">

```

```

<contents datatype="plainText"/>
</adminSpec>
<adminSpec name="securitySubset" datcatId="ISO12620A-100309">
<contents datatype="picklist">public confidential</contents>
</adminSpec>
<refSpec name="see" datcatId="ISO12620A-101801">
  <contents targetType="element" forTermComp="yes"/>
</refSpec>
<descripSpec name="sequentiallyRelatedConcept" datcatId="ISO12620A-070206">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<termNoteSpec name="shortFormFor" datcatId="ISO12620A-02013004">
  <contents datatype="basicText" targetType="term"/>
</termNoteSpec>
<adminSpec name="sortKey" datcatId="ISO12620A-100602">
  <contents datatype="plainText"/>
</adminSpec>
<adminSpec name="source" datcatId="ISO12620A-1019">
  <contents/>
</adminSpec>
<adminSpec name="sourceIdentifier" datcatId="ISO12620A-1020">
  <contents datatype="plainText" targetType="bibl"/>
</adminSpec>
<adminNoteSpec name="sourceType" datcatId="ISO12620A-102001">
  <contents datatype="picklist">parallelText backgroundText</contents>
</adminNoteSpec>
<descripSpec name="spatiallyRelatedConcept" datcatId="ISO12620A-070208">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="subjectField" datcatId="ISO12620A-04">
  <contents datatype="plainText"/>
  <levels>termEntry</levels>
</descripSpec>
<descripSpec name="subordinateConceptGeneric" datcatId="ISO12620A-07020301">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="subordinateConceptPartitive" datcatId="ISO12620A-07020302">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<hiSpec name="subscript" datcatId="">
  <contents/>
</hiSpec>
<adminSpec name="subsetOwner" datcatId="ISO12620A-10020210">
  <contents datatype="plainText"/>
</adminSpec>
<descripSpec name="superordinateConceptGeneric" datcatId="ISO12620A-07020202">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<descripSpec name="superordinateConceptPartitive" datcatId="ISO12620A-07020203">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<hiSpec name="superscript" datcatId="">
  <contents/>
</hiSpec>
<termCompListSpec name="syllabification" datcatId="ISO12620A-0206">
  <contents forTermComp="yes"/>
</termCompListSpec>

```

```

<descripSpec name="table" datcatId="ISO12620A-050504">
  <contents datatype="plainText" targetType="binaryData"/>
  <levels>langSet termEntry term</levels>
</descripSpec>
<termNoteSpec name="temporalQualifier" datcatId="ISO12620A-020305">
  <contents datatype="picklist">archaicTerm outdatedTerm obsoleteTerm</contents>
</termNoteSpec>
<descripSpec name="temporallyRelatedConcept" datcatId="ISO12620A-070207">
  <contents datatype="basicText" targetType="entry"/>
  <levels>langSet termEntry</levels>
</descripSpec>
<termCompListSpec name="termElement" datcatId="ISO12620A-020802">
  <contents forTermComp="yes"/>
</termCompListSpec>
<termNoteSpec name="termLocation" datcatId="n/a">
  <contents datatype="plainText"/>
</termNoteSpec>
<termNoteSpec name="termProvenance" datcatId="ISO12620A-020401">
  <contents datatype="picklist">transdisciplinaryBorrowing translanguagualBorrowing
loanTranslation neologism</contents>
</termNoteSpec>
<termNoteSpec name="termStructure" datcatId="ISO12620A-020804">
  <contents datatype="plainText" forTermComp="yes"/>
</termNoteSpec>
<termNoteSpec name="termType" datcatId="ISO12620A-0201">
  <contents datatype="picklist">abbreviation acronym clippedTerm commonName entryTerm
equation formula fullForm initialism internationalScientificTerm internationalism
logicalExpression phraseologicalUnit partNumber shortForm shortcut sku symbol synonym
synonymousPhrase standardText string transliteratedForm transcribedForm variant
</contents>
</termNoteSpec>
  <descripSpec name="thesaurusDescriptor" datcatId="ISO12620A-0902">
    <contents datatype="plainText" targetType="thesaurusDescrip"/>
    <levels>termEntry</levels>
  </descripSpec>
<termNoteSpec name="timeRestriction" datcatId="ISO12620A-020306">
  <contents datatype="plainText"/>
</termNoteSpec>
<transacSpec name="transactionType" datcatId="ISO12620A-1001">
  <contents datatype="picklist">origination input modification check approval
withdrawal standardization exportation importation proposal userAccess</contents>
</transacSpec>
<termNoteSpec name="transferComment" datcatId="ISO12620A-0305">
  <contents targetType="term"/>
</termNoteSpec>
<descripSpec name="unit" datcatId="ISO12620A-0506">
  <contents datatype="plainText"/>
  <levels>term</levels>
</descripSpec>
<transacNoteSpec name="usageCount" datcatId="ISO12620A-100203">
  <contents/>
</transacNoteSpec>
<termNoteSpec name="usageNote" datcatId="ISO12620A-020301">
  <contents/>
</termNoteSpec>
<descripSpec name="video" datcatId="ISO12620A-050503">
  <contents datatype="plainText" targetType="binaryData"/>
  <levels>langSet termEntry term</levels>
</descripSpec>
<xrefSpec name="xAudio" datcatId="ISO12620A-050502">
  <contents targetType="external"/>
</xrefSpec>
<xrefSpec name="xGraphic" datcatId="ISO12620A-050501">

```

```

    <contents targetType="external"/>
  </xrefSpec>
  <xrefSpec name="xMathML" datcatId="">
    <contents targetType="external"/>
  </xrefSpec>
  <xrefSpec name="xSource" datcatId="ISO12620A-1019">
    <contents targetType="external"/>
  </xrefSpec>
  <xrefSpec name="xVideo" datcatId="ISO12620A-050503">
    <contents targetType="external"/>
  </xrefSpec>
</datCatSet>
<refObjectDefSet>
  <refObjectDef>
    <refObjectType>binaryData</refObjectType>
    <itemSpecSet type="validItemType">
      <itemSpec type="validItemType">format</itemSpec>
      <itemSpec type="validItemType">fileName</itemSpec>
      <itemSpec type="validItemType">fileDateTime</itemSpec>
      <itemSpec type="validItemType">fileAttributes</itemSpec>
      <itemSpec type="validItemType">fileOSName</itemSpec>
      <itemSpec type="validItemType">fileOSNumber</itemSpec>
      <itemSpec type="validItemType">codePage</itemSpec>
      <itemSpec type="validItemType">fileSize</itemSpec>
      <itemSpec type="validItemType">data</itemSpec>
      <itemSpec type="validItemType">description</itemSpec>
    </itemSpecSet>
  </refObjectDef>
</refObjectDefSet>
</TBXXCS>

```

STANDARDSISO.COM : Click to view the full PDF of ISO 30042:2008

## Annex D (Normative) Descriptions of the core structure elements and attributes and the default data-categories

### Annex D.1 General information about the descriptions

This annex contains a description of the TBX core structure that is defined in Annex A and the default meta data-categories that are defined in Annex C. Three types of XML constructs are described:

#### macros

Describes data types and groups of elements that are used in multiple element declarations. These macros are then referred to in the element declarations that inherit them. This section also corresponds to some of the entities that are declared in the core DTD.

#### attribute classes

Describes individual attributes, and groups of attributes, that are used in multiple element declarations. These attribute classes are then referred to in the element declarations that inherit them. If an attribute is only used by one element, it is described at the location that the element is described. This section also corresponds to some of the entities that are declared in the core DTD.

#### elements

Describes the TBX elements and their attributes.

**NOTE** Values of the *type* attribute that are used in the default TBX TML are described in Annex C.5.

The following table describes the parts of the descriptions and how to read them.

<b>Members</b>	Indicates which elements and attribute classes use (inherit) the attribute or attribute class described.
<b>Attributes</b>	Indicates the attributes and attribute classes that are permissible for the element or attribute class that is currently being defined. A link to the attribute class declaration is followed by the names of the attributes in that class (indicated after the @ symbol, which is the Relax NG representation) in parentheses. If the attribute is unique to the element or attribute class that is currently defined its declaration is provided directly. The attribute is either optional or required. NOTE: An the element's content may be constrained to a picklist when the element has certain values of its <i>type</i> attribute. The meanings of the picklist values are not defined in this International Standard. To obtain information about the meaning of picklist values, refer to ISO 12620.
<b>Used by</b>	Indicates which elements or element groups (macros) can contain (i.e. be the parent of) the element described.
<b>May contain</b>	Indicates which elements or data types can occur in the element described. This row indicates the default content model of the meta data-category. For instance, in the case of <descrip>, this row indicates that it can take the child elements <hi>, <bpt>, <ept>, and so forth. However, the default content model may be further constrained by the value of the <i>type</i> attribute of the element. For instance, when the <descrip> element has a <i>type</i> attribute value of 'range', none of the aforementioned child elements are permitted. Any such additional constraints are indicated in the description of the data-category. <b>Note:</b> This section does not indicate which elements are mandatory and which are optional. For this information, refer to the DTD (Annex A) or the integrated RelaxNG schema referred to in Annex F.

### D.2 Macros

#### entity.auxInfo

The auxInfo entity corresponds to information that can be associated with any one of three levels: the terminological entry level (<termEntry> (i.e., the concept level), the language level (<langSet>), and the term level (<ntig> or its simplified version <tig>).

**entity.basicText**

Basic text is used for terms and term-like elements as well as a few other elements. It allows plainText plus <hi> for some limited embedded markup.

**entity.noteText**

Note text, which is used in definitions and contextual examples and similar elements, allows more inline markup than basicText.

**entity.noteLinkInfo**

This entity groups elements for administrative information, notes, and links. It is used by the following grouping elements: <termCompGrp>, <termNoteGrp> and <itemGrp>.

**D.3 Attribute classes****id**

The id attribute class corresponds to /elementIdentifier/ in ISO 12620:1999. It indicates a unique identifier of an element. Element identifiers can include entry identifiers (eid-...), concept identifiers (cid-...), term identifiers (tid-...), among others. This attribute is optional.

**Members** IDLangTgtDtyp [ admin adminNote descrip descripNote ref termNote transac transacNote] idlang [ change note revisionDesc sourceDesc termComp title titleStmt] idtype [ item itemSet] adminGrp back body date descripGrp encodingDesc fileDesc foreign itemGrp martifHeader ntig p publicationStmt refObject refObjectList term termCompGrp termCompList termEntry termGrp termNoteGrp text tig transacGrp xref

**Attributes** @id  
**Status** Optional

**Datatype** xsd:ID

**idlang**

The idlang attribute class combines the *id* attribute and the *xml:lang* attribute. Both are optional for elements that use this attribute class.

**Members** [change](#) [note](#) [revisionDesc](#) [sourceDesc](#) [termComp](#) [title](#) [titleStmt](#)

**Attributes** [id](#) (@id) [lang](#) (@xml:lang)

**IDLangTgtDtyp**

The IDLangTgtDtyp attribute class groups together the attributes *id*, *xml:lang*, *target* and *datatype*.

**Members** [admin](#) [adminNote](#) [descrip](#) [descripNote](#) [ref](#) [termNote](#) [transac](#) [transacNote](#)

**Attributes** [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)

@datatype

The data type of plainText, or what is known as PCDATA.

**Status** Optional

**Datatype** text

**idtype**

The idtype attribute class combines the *id* attribute and the *type* attribute. Both are optional for elements that use this attribute class.

**Members** [item](#) [itemSet](#)

**Attributes** [id](#) (@id)

@type

Indicates the type of the meta data-category.

**Status** Optional

**Datatype** text

### lang

The lang attribute class corresponds to the attribute *xml:lang*.

**Members** IDLangTgtDtyp [ admin adminNote descrip descripNote ref termNote transac transacNote] idlang [ change note revisionDesc sourceDesc termComp title titleStmnt] foreign hi p

**Attributes** @xml:lang

Indicates the language of the element content. The allowed values are found in IETF RFC 4646 or its successor. This attribute is optional for elements that use this attribute class. For a description, see the information at the <marfif> element.

**Status** Optional

**Datatype** text

### target-IDREF

The target-IDREF attribute class corresponds to the *target* attribute, which is a pointer to another element. The link is established when the value of the *target* attribute of the *source* element is the same as the value of the *xml:id* attribute of the target element. When this attribute is used with <termNote>, the target is often another term, such as in the tag <termNote type='abbreviatedFormFor' target='tid-1234'>. It is also available for the data-category transferComment to indicate the term being commented about. The pointer is a unique identifier. The content of elements that take this attribute is typically the display value of the target element, such as the target term. The elements that take this attribute cannot be empty; display content is mandatory. In respect of the principle of term autonomy, any term that is contained by this element should have its own entry in the data collection.

**Members** IDLangTgtDtyp [ admin adminNote descrip descripNote ref termNote transac transacNote] hi

**Attributes** @target

**Status** Optional

**Datatype** xsd:IDREF

## D.4 Elements

### <admin>

Contains information of an administrative nature for the node (parent element) in question, such as the source of information, or the project or client for which it applies. The type of administrative information is indicated by the value of the *type* attribute. It can appear alone, or, if additional information needs to be provided, such as a note or a reference, it can be nested in an <adminGrp> element.

NOTE This element can occur at any of the three levels of the terminological entry. Several types, however, logically should occur at restricted levels, as indicated with the corresponding data-category descriptions. For instance, databaseType and conceptOrigin would typically occur at the termEntry level. Care should be taken to insert the <admin> element at the appropriate level of the entry according to its *type* attribute value and the parent node to which the administrative information applies.

**Attributes** [IDLangTgtDtyp](#) [(@datatype) [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)]

@type

Indicates the data-category type of the <admin> element, such as /originatingDatabase/ or /entrySource/.

**Status** Required

**Used by** [adminGrp](#) [descripGrp](#) [entity.auxInfo](#) [entity.noteLinkInfo](#)

**May contain** [bpt](#) [ept](#) [foreign](#) [hi](#) [ph](#)

### <adminGrp>

An element that contains a group of elements that contain administrative information.

**Attributes** [id](#) (@id)

**Used by** [descripGrp](#) [entity.auxInfo](#) [entity.noteLinkInfo](#)

**May contain** [admin](#) [adminNote](#) [note](#) [ref](#) [xref](#)

### <adminNote>

An administrative note about some data. Currently this element is only used to specify the type of source text from which a term was found, and therefore only one value is currently specified for the *type* attribute. However, additional *type* attribute values can be defined for additional purposes in a user-specific XCS file.

**Attributes** [IDLingTgtDtyp](#) [(@datatype) [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)]  
@type

**Status** Required

**Used by** [adminGrp](#)

**May contain** Character data only

### <back>

The root element of the back matter of a TBX document instance.

**Attributes** [id](#) (@id)

**Used by** [text](#)

**May contain** [refObjectList](#)

### <body>

A nesting element that contains terminological entries (<termEntry>).

**Attributes** [id](#) (@id)

**Used by** [text](#)

**May contain** [termEntry](#)

### <bpt>

The <bpt> tag encloses the beginning tag of a tag pair.

**Attributes** @i

**Status** Optional

**Datatype** text

## ISO 30042:2008(E)

@type	
Status	Optional
Datatype	text

Used by [entity.noteText](#)

May contain Character data only

### <change>

Information about one particular change event in the revision history of a TBX document instance.

Attributes [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

Used by [revisionDesc](#)

May contain [p](#)

### <date>

A date in ISO format. Permissible date values comply with ISO 8601 and shall use the format yyyy-mm-dd. The date element is used with terminology management data-categories in a <transacGrp> element.

Attributes [id](#) (@id)

Used by [transacGrp](#)

May contain Character data only

### <descrip>

An element that contains descriptive information about a concept, or relations to other concepts. The type of information that the element contains, and any restrictions on the permissible values of the element, are determined by the value of the *type* attribute.

Attributes [IDLingTgtDtyp](#) [([datatype](#)) [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)]  
@type

Indicates the data-category type of the <descrip> element, such as definition or associated-concept.

Status	Required
--------	----------

Used by [descripGrp](#) [entity.auxInfo](#)

May contain [bpt](#) [ept](#) [foreign](#) [hi](#) [ph](#)

### <descripGrp>

Contains one <descrip> element as well as additional child elements for associated administrative information.

Attributes [id](#) (@id)

Used by [entity.auxInfo](#)

May contain [admin](#) [adminGrp](#) [descrip](#) [descripNote](#) [note](#) [ref](#) [transacGrp](#) [xref](#)

### <descripNote>

Element used to indicate the type of definition or context with which it is associated.

**Attributes** [IDLangTgtDtyp](#) [(@datatype) [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)]  
**@type**  
 Indicates the data-category type of the <descripNote> element, such as 'contextType'.  
**Status** Required

**Used by** [descripGrp](#)

**May contain** Character data only

### <encodingDesc>

Includes a <p> element that itself includes a pointer to the XCS file, or an embedded XCS file, and, if the TBX file is not in Unicode, any information about the encoding.

**Attributes** [id](#) (@id)

**Used by** [martifHeader](#)

**May contain** [p](#)

### <ept>

The <ept> tag encloses the beginning tag of a tag pair.

**Attributes** [@i](#)  
**Status** Optional  
**Datatype** text

**Used by** [entity.noteText](#)

**May contain** Character data only

### <fileDesc>

A nesting element containing child elements that describe the TBX document instance.

**Attributes** [id](#) (@id)

**Used by** [martifHeader](#)

**May contain** [publicationStmnt](#) [sourceDesc](#) [titleStmnt](#)

### <foreign>

Used to mark a segment of text that is in a different language from the surrounding text.

**Attributes** [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

**Used by** [entity.noteText](#)

**May contain** [bpt](#) [ept](#) [foreign](#) [hi](#) [ph](#)

### <hi>

Highlights a segment of text and optionally points to another element.

**Attributes** [target-IDREF](#) [(@target) [lang](#) (@xml:lang)]  
**@type**

## ISO 30042:2008(E)

**Status** Optional

**Used by** [entity.basicText](#) [entity.noteText](#)

**May contain** Character data only

### <item>

One descriptor of an object in the back matter of a TBX document instance.

**Attributes** [idtype](#) [([@type](#)) [id](#) ([@id](#))]

**Used by** [itemGrp](#) [itemSet](#) [refObject](#)

**May contain** [bpt](#) [ept](#) [foreign](#) [hi](#) [ph](#)

### <itemGrp>

Contains one <item> element as well as additional child elements for associated administrative information.

**Attributes** [id](#) ([@id](#))

**Used by** [itemSet](#) [refObject](#)

**May contain** [admin](#) [adminGrp](#) [item](#) [note](#) [ref](#) [transacGrp](#) [xref](#)

### <itemSet>

A list of <item> or <itemGrp>.

**Attributes** [idtype](#) [([@type](#)) [id](#) ([@id](#))]

**Used by** [refObject](#)

**May contain** [item](#) [itemGrp](#)

### <langSet>

A nesting element that contains all the information in a terminological entry pertaining to one language, including all the <tig> or <ntig> elements (terms and associated information) for that language.

**Attributes** [@xml:lang](#)

Indicates the language of the language section. This attribute is required for the <langSet> element. See also the description on the <martif> element.

**Status** Required

**Datatype** text

**Used by** [termEntry](#)

**May contain** [admin](#) [adminGrp](#) [descrip](#) [descripGrp](#) [note](#) [ntig](#) [ref](#) [tig](#) [transacGrp](#) [xref](#)

### <martif>

The root element of a TBX document instance.

**Attributes** [@xml:lang](#)

Indicates the default language of the entire <martif> element content. The allowed values of the *xml:lang* attribute are found in IETF RFC 4646 or its successor, as indicated in IETF BCP 47. This attribute is required for the <martif>

element.

**Status** Required  
**Datatype** text

NOTE - The value of the *xml:lang* attribute inherits downward through the XML document unless overridden by another *xml:lang* attribute. The language specified in the <martif> element becomes the working language of the entire TBX file. Each <langSet> element shall also specify a language that applies to that language section. This becomes the object language of the language section. Thus, a definition at the terminological entry level is assumed to be in the working language of the TBX file unless otherwise specified, and a note in a language section is assumed to be in the same language as the terms in that language section unless otherwise specified.

@type

Indicates the type of martif document. If the *type* attribute value 'TBX' is used, then the TBX document instance shall adhere to the default XCS file. If the TBX document instance adheres to another XCS file, then the *type* attribute value can consist of 'TBX-' plus another word associated with the XCS file, such as 'TBX-Basic'.

**Status** Required  
**Datatype** text

**Used by** n/a

**May contain** [martifHeader](#) [text](#)

### <martifHeader>

A grouping element that contains child elements which describe the TBX document instance.

**Attributes** [id](#) (@id)

**Used by** [martif](#)

**May contain** [encodingDesc](#) [fileDesc](#) [revisionDesc](#)

### <note>

Any kind of note.

**Attributes** [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

**Used by** [adminGrp](#) [descripGrp](#) [titleStm](#) [transacGrp](#) [entity.auxInfo](#) [entity.noteLinkInfo](#)

**May contain** [bpt](#) [ept](#) [foreign](#) [hi](#) [ph](#)

### <ntig>

A *nesting term information group*. A grouping element that contains child elements describing a term. It is equivalent to a <tig> except that it allows the description of term components.

**Attributes** [id](#) (@id)

**Used by** [langSet](#)

**May contain** [admin](#) [adminGrp](#) [descrip](#) [descripGrp](#) [note](#) [ref](#) [termGrp](#) [transacGrp](#) [xref](#)

### <p>

Paragraph text that is contained by the <martifHeader>. By using a *type* attribute, it can also contain essential information about the content or location of the XCS file.

**Attributes** [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

@type

**Status** Optional

**Only when used in encodingDesc, the values are:**

- XCSURI - The URI of the XCS file that constrains the content of the TBX document instance.
- DCSName - The name of the XCS file that constrains the content of the TBX document instance. This value is provided for backwards compatibility purposes and is deprecated; use XCSURI instead.
- XCSContent - The content of an XCS file, embedded directly in the TBX document instance. Note that this content cannot be validated by the TBX DTD and therefore should be removed from the TBX document instance, or commented out, prior to such validation.

**Used by** [change encodingDesc](#) [publicationStmnt](#) [sourceDesc](#)

**May contain** [bpt](#) [ept](#) [foreign](#) [hi](#) [ph](#)

**<ph>**

Placeholder - Used to encapsulate markup codes that do not occur in logical pairs with other codes. Examples include the XHTML `<br/>` and `<hr/>` tags. This element should not be used to encapsulate codes that occur in pairs (such as the XHTML `<em>` and `</em>` tags), which should instead be encapsulated with `<bpt>` and `<ept>`.

**Attributes** @type  
**Status** Optional  
**Datatype** text

**Used by** [entity.noteText](#)

**May contain** Character data only

**<publicationStmnt>**

Any information pertaining to the publication of the TBX document instance.

**Attributes** [id](#) (@id)

**Used by** [fileDesc](#)

**May contain** [p](#)

**<ref>**

A cross-reference that points to another element within the `<martif>` element of the TBX document instance. The content of the element is display text of the target element (i.e. the term), and shall be in plainText.

**Attributes** [IDLangTgtDtyp](#) [(@datatype) [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)]  
 @type

**Status** Required

**Used by** [adminGrp](#) [descripGrp](#) [transacGrp](#) [entity.auxInfo](#) [entity.noteLinkInfo](#)

**May contain** Character data only

**<refObject>**

An element in the back matter of a TBX document instance that contains a series of items that describe one referenced object.

**Attributes** [id](#) (@id)

**Used by** [refObjectList](#)

**May contain** [item](#) [itemGrp](#) [itemSet](#)

**<refObjectList>**

A list of objects that are referenced in the back matter of a TBX document instance.

**Attributes** [id](#) (@id)

[@type](#)

**Status** Required

**Datatype** text

**Used by** [back](#)

**May contain** [refObject](#)

**<revisionDesc>**

Information about any revisions that have been carried out on the content of the TBX document instance or on the source database from which it is derived.

**Attributes** [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

**Used by** [mariffHeader](#)

**May contain** [change](#)

**<sourceDesc>**

Any information about the source of the TBX document instance.

**Attributes** [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

**Used by** [fileDesc](#)

**May contain** [p](#)

**<term>**

The term that is being described in a <tig> or <ntig>. This element, as well as other term-like elements such as those mentioned in the table [Types of terms, relations to terms](#) in section 9, can contain a <hi> element to allow a limited amount of inline markup. This is intended to handle markup requirements in special cases such as may be required for terms that represent scientific concepts. However, it is strongly recommended to use inline markup only when necessary to represent the term in its base form. Do not use the <hi> element for presentational styles chosen for esthetic purposes.

**Attributes** [id](#) (@id)

**Used by** [termGrp](#) [tig](#)

**May contain** [hi](#)

**<termComp>**

A term component. It includes the text of a part of a word (such as a morpheme or a syllable) or one of the words making up a multi-word term. The type of term component will depend on the value of the *type* attribute of the <termCompList> element.

**Attributes** [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

**Used by** [termCompGrp](#) [termCompList](#)

## ISO 30042:2008(E)

**May contain** Character data only

### <termCompGrp>

A grouping element that includes one <termComp> plus additional elements describing the term component, such as notes, administrative information, and cross references.

**Attributes** [id](#) (@id)

**Used by** [termCompList](#)

**May contain** [admin](#) [adminGrp](#) [note](#) [ref](#) [termComp](#) [termNote](#) [termNoteGrp](#) [transacGrp](#) [xref](#)

### <termCompList>

A grouping element that contains at least one <termComp> element (usually more) or at least one <termCompGrp> element plus optionally some additional information about the term components. This element is designed to contain the components of a term. The value of the *type* attribute indicates what type of term components it contains.

**Attributes** [id](#) (@id)

[@type](#)

Indicates the type of the <termCompList>, such as hyphenation. The value indicates on what basis the term components were identified, such as by hyphenation parts or by syllables.

**Status** Required

**Used by** [termGrp](#)

**May contain** [admin](#) [adminGrp](#) [descrip](#) [descripGrp](#) [note](#) [ref](#) [termComp](#) [termCompGrp](#) [transacGrp](#) [xref](#)

### <termEntry>

The root element of a terminological entry. It shall contain at least one language section.

**Attributes** [id](#) (@id)

**Used by** [body](#)

**May contain** [admin](#) [adminGrp](#) [descrip](#) [descripGrp](#) [langSet](#) [note](#) [ref](#) [transacGrp](#) [xref](#)

### <termGrp>

Element used in an <ntig> which contains one <term> and optionally additional information about the term. This element is primarily used in conjunction with <ntig> when the term is broken down into components.

**Attributes** [id](#) (@id)

**Used by** [ntig](#)

**May contain** [term](#) [termCompList](#) [termNote](#) [termNoteGrp](#)

### <termNote>

A meta data-category used for describing terms. A *type* attribute specifies what kind of information is included in a particular instance of this element.

**Attributes** [IDLangTgtDtyp](#) [(@datatype) [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)]

[@type](#)

Indicates the data-category type of the <termNote>, such as partOfSpeech or termType.

**Status** Required

**Used by** [termCompGrp](#) [termGrp](#) [termNoteGrp](#) [tig](#)

**May contain** [bpt](#) [ept](#) [foreign](#) [hi](#) [ph](#)

### <termNoteGrp>

A grouping element that contains one <termNote> element, and [auxiliary information](#) such as administrative information, transaction information, notes and cross references. It does not allow any <descrip> elements. It can only appear at the term (tig or ntig) level and below.

**Attributes** [id](#) (@id)

**Used by** [termCompGrp](#) [termGrp](#)

**May contain** [admin](#) [adminGrp](#) [note](#) [ref](#) [termNote](#) [transacGrp](#) [xref](#)

### <text>

The element of a TBX document instance that occurs after the [martif](#) element. This element is provided to maintain compatibility with the TEI.

**Attributes** [id](#) (@id)

**Used by** [martif](#)

**May contain** [back](#) [body](#)

### <tig>

A *term information group*. A grouping element that contains child elements describing a term. It is often referred to as the *term section*. If the section needs to include information about the components of a term (such as morphology) or the individual words of a multi-word term, the equivalent element <ntig> should be used which allows deeper nesting.

**Attributes** [id](#) (@id)

**Used by** [langSet](#)

**May contain** [admin](#) [adminGrp](#) [descrip](#) [descripGrp](#) [note](#) [ref](#) [term](#) [termNote](#) [transacGrp](#) [xref](#)

### <title>

The title of the TBX document instance.

**Attributes** [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

**Used by** [header](#) [titleStmt](#)

**May contain** Character data only

### <titleStmt>

A nesting element containing the title and any notes about the TBX document instance.

**Attributes** [idlang](#) [[id](#) (@id) [lang](#) (@xml:lang)]

## ISO 30042:2008(E)

Used by [fileDesc](#)

May contain [note title](#)

### <transac>

A record that indicates the stage of the entry within the overall process of creation, approval, and use of a terminology entry. Currently this element is only used to specify the type of transaction, and therefore only one value is currently specified for the *type* attribute. However, additional *type* attribute values can be defined for additional purposes in a user-specific XCS file.

**Attributes** [IDLingTgtDtyp](#) [(@datatype) [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)]

[@type](#)

The attribute value *transactionType* ensures that the content of the <transac> element will be a transaction type.

**Status** Required

Used by [transacGrp](#)

May contain Character data only

### <transacGrp>

An element that includes child elements that provide information about a transaction.

**Attributes** [id](#) (@id)

Used by [descripGrp](#) [entity.auxInfo](#) [entity.noteLinkInfo](#)

May contain [date](#) [note](#) [ref](#) [transac](#) [transacNote](#) [xref](#)

### <transacNote>

Information about a transaction, such as the name of the person who performed the transaction. The content of this element shall be plainText.

**Attributes** [IDLingTgtDtyp](#) [(@datatype) [lang](#) (@xml:lang) [id](#) (@id) [target-IDREF](#) (@target)]

[@type](#)

**Status** Required

Used by [transacGrp](#)

May contain Character data only

### <xref>

A linking element that points to an external object using a URI (a URL or other Web address). The *target* attribute is required. The content of the element is display text representing the target object, such as a title, file name, or short description.

**Attributes** [id](#) (@id)

[@target](#)

The URI of the external object.

**Status** Required

**Datatype** text

NOTE Unlike the *target* attributes on most TBX elements, which point to other elements in the TBX XML document instance (hence, they have an IDREF datatype), the *target* attribute of an <xref> element points outside the TBX document, presumably to objects identified by URIs. Therefore, the datatype of the *target* attribute for an <xref> element is specified as CDATA.

@type  
**Status** Optional

**Used by** [adminGrp](#) [descripGrp](#) [transacGrp](#) [entity.auxInfo](#) [entity.noteLinkInfo](#)

**May contain** Character data only

## D.5 Default data-categories

The data-category definitions in this section are derived from ISO 12620. They are grouped according to the meta data-category with which they are instantiated through the type value of the meta data-category. For example, for the tag `<descrip type='definition'>`, the data-category `/definition/` is described in the **descrip** section.

### admin

#### annotatedNote

A note that can be coupled with an `<adminNote>` element to indicate the source of the note or other administrative information about the note. If no source or additional information about the note is required, use the `<note>` element.

#### applicationSubset

An identifier that indicates that the parent element to which it applies is associated with a specific application. The content shall be in plainText.

#### businessUnitSubset

An identifier that indicates that the parent element to which it applies is associated with a specific department, division, or other unit of an enterprise. The content shall be in plainText.

#### conceptOrigin

The institution, region, community, culture, country, or other venue from which a concept covered by a terminological entry originates. Because this data-category provides information about the origin of the concept, it should occur at the `<termEntry>` level. The content shall be in plainText.

#### customerSubset

An identifier that indicates that the parent element to which it applies is associated with a specific customer. The content shall be in plainText.

#### databaseType

An identifier that indicates the type of database from which the parent element originated. This element should normally occur at the `<termEntry>` level because typically entire entries originate from or are imported from another database. The content shall be in plainText.

#### domainExpert

Subject field specialist who provides information for a terminological entry, such as a definition or target language equivalent. An `<admin>` element that has a *type* attribute value of 'domainExpert' can also have a *target* attribute pointing to the identifier of a reference entry in the back matter of the TBX document instance that provides information about the domain expert. The content shall be in plainText.

#### elementWorkingStatus

A value indicating the level of completeness and accuracy of an element (field, record, entry) within a terminological collection. The element `<admin>` when its *type* attribute has the value 'elementWorkingStatus' shall contain one of the following values:

- starterElement
- workingElement
- consolidatedElement
- archiveElement
- importedElement
- exportedElement

#### entrySource

A database or file from which the entry was imported. The content shall be in plainText.

**environmentSubset**

An identifier that indicates a specific computer environment that the parent element is associated with. The content shall be in plainText.

**indexHeading**

A term chosen to be used in an index to represent an item or a concept in a document. The content shall be in plainText.

**keyword**

A word or group of words, possibly in lexicographically standardized (lemmatized) form, taken out of the text of a document, which characterizes the text and enables its retrieval. For example, the terms data-category and terminological entry are keywords for this international standard. In a terminological database that contains the terminology of ISO 30042, they could be identified as keywords by using this element. The content shall be in plainText.

**originatingDatabase**

The database from which the parent element originated. The content shall be in plainText.

**originatingInstitution**

An institution (i.e., company, government agency, etc.) that provided the information in the parent element. The content shall be in plainText.

**originatingPerson**

An individual who provided the information in the parent element. The content shall be in plainText.

**productSubset**

An identifier assigned to a term, record, or entry to indicate that it is related to or used in a product. The content shall be in plainText, and preferably one of a predefined set of picklist values.

**projectSubset**

An identifier assigned to a term, record, or entry to indicate that it is related to or used in a project. The content shall be in plainText, and preferably one of a predefined set of picklist values.

**searchTerm**

A term entered in a term entry for purposes of retrieval. The content shall be in plainText

**securitySubset**

A classification of a term, record, or entry indicating that its distribution is restricted for security reasons. Sometimes, terms are not released to the public until the material that uses the term (product, service, etc.) is released to the public. Until that time, this element serves to protect the data from distribution. The element <admin> when its *type* attribute has the value 'securitySubset' shall contain one of the following values:

- public
- confidential

**sortKey**

A character string used for sorting and merging operations. A sort key can enable alphabetic or systematic access. The content shall be in plainText.

**source**

A complete citation of the source of the information in the parent element.

**sourceIdentifier**

A pointer to a bibliographic entry that contains information about the source of the parent element. An <admin> element that has a *type* attribute value of 'sourceIdentifier' also has a *target* attribute containing the identifier of the bibliographic entry. The content of the element is a display value for the source.

**subsetOwner**

The owner of the subset of terminological entries to which this entry belongs. The content shall be in plainText.

**adminNote****sourceType**

A property describing the source text, i.e. whether it is actual parallel text (text that is equivalent to text in another language) or background material. The content of the <adminNote> element when it has a *type* attribute value of 'sourceType' shall be one of the following:

- parallelText
- backgroundText

**noteSource**

The source of a note that is expressed in a <admin> element having the *type* attribute value 'annotatedNote'.

**descrip****antonymConcept**

A concept that is opposite in meaning to the concept represented by the entry containing the <descrip> element. A <descrip> element that has a *type* attribute value of 'antonymConcept' can also have a *target* attribute pointing to the entry ID of the <termEntry> that contains the antonym concept. The display term in the content of the element shall be expressed in basicText.

**associatedConcept**

A concept that exhibits a relation to the subject concept that can be established on the basis of non-hierarchical thematic connections. A <descrip> element that has a *type* attribute value of 'associatedConcept' can also have a *target* attribute pointing to the entry ID of the <termEntry> that contains the associated concept. For example, teacher and school represent associated concepts.

**audio**

A name or other identifier of an audio file: sound, spoken words, music, or other audible representations used to illustrate or explain terms or concepts. A <descrip> element that has a *type* attribute value of 'audio' can also have a *target* attribute pointing to the ID of a <refObject> in the back matter of the TBX document instance that describes and points to the audio file. The element content shall be in plainText. NOTE If a direct pointer to the audio file is preferred within the <termEntry>, rather than pointing to the back matter, use an <xref> instead.

**broaderConceptGeneric**

A concept two or more levels of abstraction higher than subject concept in a generic hierarchical concept system. A <descrip> element that has a *type* attribute value of 'broaderConceptGeneric' can also have a *target* attribute pointing to the entry ID of the <termEntry> that contains the broader concept. Generic relations between concepts are "is-a" relations. This means that the narrower concept "is an" instance of its broader generic concept. For example, a dog is a canine animal, and a canine animal is a mammal. Therefore, mammal is a broader generic concept for dog.

**broaderConceptPartitive**

A concept two or more levels of abstraction higher than subject concept in a part-whole or meronymic hierarchical concept system. A <descrip> element that has a *type* attribute value of 'broaderConceptPartitive' can also have a *target* attribute pointing to the entry ID of the <termEntry> that contains the broader concept. For example, a floor board is part of a house (broader partitive concept), but it can be more narrowly defined as being part of the floor (superordinate partitive concept), which in turn is part of the house. Partitive relations between concepts are "has-a" relations. This means that the broader partitive concept "has" its narrower concepts as parts of itself.

**characteristic**

An abstraction of a property of an object or of a set of objects. For example, *flammability* is a characteristic of fuel. Essential characteristics can be used to define concepts. The element content shall be in plainText.

**classificationCode**

A set of symbols, with rules for their application, used to represent classes and their interrelations. A <descrip>

element that has a *type* attribute value of 'classificationCode' can also have a *target* attribute pointing to the ID of a bibliographic <refobject> in the back matter describing the classification code. The element content shall be in plainText. NOTE If a direct pointer to the classification code is preferred within the <termEntry>, rather than pointing to the back matter, use an <xref> instead.

#### **conceptPosition**

The position of a concept in a concept system. A <descrip> element that has a *type* attribute value of 'conceptPosition' can also have a *target* attribute pointing to the ID of a position in the concept system. The element content shall be in plainText.

#### **context**

A text which illustrates a concept or a term, by containing the concept designation itself. Contexts must be authentic, that is, they must be obtained from an existing source, and not created by the terminologist. Contexts are documented very frequently in terminology collections. Contexts can provide information for determining term usage and collocations. In TBX, the context as a term-related data-category. A context can be further categorized according to context type.

#### **coordinateConceptGeneric**

A concept having the same nearest superordinate concept as the concept being described by this <descrip> element in a generic concept system. A <descrip> element that has a *type* attribute value of 'coordinateConceptGeneric' can also have a *target* attribute pointing to the ID of the coordinate concept.

#### **coordinateConceptPartitive**

A concept having the same nearest superordinate concept as the concept being described by this <descrip> element in a partitive concept system. A <descrip> element that has a *type* attribute value of 'coordinateConceptPartitive' can also have a *target* attribute pointing to the ID of the coordinate concept.

#### **definition**

A representation of a concept by a descriptive statement which serves to differentiate that concept from related concepts.

#### **example**

A text which illustrates a concept or a term, by providing an example of the object designated by the concept or term, and not necessarily containing the concept designation itself. Some databases use *example* as a label for contextual references. The content of those data-categories should be converted to the data-category [context](#) for interchange purposes.

#### **explanation**

A statement that describes and clarifies a concept and makes it understandable, but does not necessarily differentiate it from other concepts.

#### **figure**

A name or other description of a diagram, picture, or other graphic material used to illustrate a concept or group of concepts. A <descrip> element that has a *type* attribute value of 'figure' can also have a *target* attribute pointing to the ID of a <refObject> in the back matter of the TBX document instance that describes and points to the file that contains the figure. The element content shall be in plainText.

NOTE If a direct pointer to the file that contains the figure is preferred within the <termEntry>, rather than pointing to the back matter, use an <xref> instead.

#### **otherBinaryData**

Any binary data not covered by the categories [figure](#), [audio](#), [video](#), and [table](#). A <descrip> element that has a *type* attribute value of 'otherBinaryData' can also have a *target* attribute pointing to the ID of a <refObject> in the back matter of the TBX document instance that describes and points to the binary file. The element content shall be in plainText. NOTE If a direct pointer to the binary file is preferred within the <termEntry>, rather than pointing to the back matter, use an <xref> instead.

#### **quantity**

A measurable quantity associated with a concept. The element content shall be in plainText.

#### **range**

The pair of limits within which a quantity is measured, as expressed by stating the lower and upper permissible

values. The element content shall be in plainText.

#### **relatedConcept**

A concept that has an associative relation to another concept, such as teacher and school. A <descrip> element that has a *type* attribute value of 'relatedConcept' can also have a *target* attribute pointing to the entry identifier of the related concept.

#### **relatedConceptBroader**

A concept that is broader at some level with respect to a related or associated concept, but without comprising any reference to a generic or partitive system. A <descrip> element that has a *type* attribute value of 'relatedConceptBroader' can also have a *target* attribute pointing to the entry identifier of the related concept.

#### **relatedConceptNarrower**

A concept that is narrower at some level with respect to a related or associated concept, but without comprising any reference to a generic or partitive system. A <descrip> element that has a *type* attribute value of 'relatedConceptNarrower' can also have a *target* attribute pointing to the entry identifier of the related concept.

#### **reliabilityCode**

A code assigned to a data-category or record indicating accuracy and/or completeness. The content of the <descrip> element when it has a *type* attribute value of 'reliabilityCode' shall be a value from 1 (least reliable) to 10 (most reliable).

#### **sampleSentence**

A sentence composed (frequently by a terminologist or language planner) to illustrate the use of a term in cases where an authentic context is unavailable. Although contextual information should ideally be taken from actual texts written in the language in question, such contextual references may not exist for newly coined or suggested terms.

#### **sequentiallyRelatedConcept**

A concept that has a relation to the subject concept based on spatial or temporal proximity. Temporal relations and spatial relations can be sequential. A <descrip> element that has a *type* attribute value of 'sequentiallyRelatedConcept' can also have a *target* attribute pointing to the entry identifier of the related concept.

#### **spatiallyRelatedConcept**

A concept that has a relation of dependence with the subject concept, referring to objects based on their relative physical positions. A <descrip> element that has a *type* attribute value of 'spatiallyRelatedConcept' can also have a *target* attribute pointing to the entry identifier of the related concept.

#### **subjectField**

A field of special knowledge. Subject fields can be expressed in multiple levels, for example: Subject field (level 1): disease, Subject field (level 2): cancer, Subject field (level 3): non-Hodgkin's lymphoma. Subject fields shall be expressed in plainText, and preferably be selected from a defined set of picklist values. Subject fields shall be specified at the concept (termEntry) level.

#### **subordinateConceptGeneric**

A concept that has a generic relation with the subject concept, and a broader intension. A <descrip> element that has a *type* attribute value of 'subordinateConceptGeneric' can also have a *target* attribute pointing to the entry identifier of the subordinate concept.

#### **subordinateConceptPartitive**

A concept that has a partitive relation with the subject concept, and represents one of the parts of the subject concept. A <descrip> element that has a *type* attribute value of 'subordinateConceptPartitive' can also have a *target* attribute pointing to the entry identifier of the subordinate concept.

#### **superordinateConceptGeneric**

A concept that has a generic relation with the subject concept, and a narrower intension. A <descrip> element that has a *type* attribute value of 'superordinateConceptGeneric' can also have a *target* attribute pointing to the entry identifier of the superordinate concept.

#### **superordinateConceptPartitive**

A concept that has a partitive relation with the subject concept, and represents the sum of the parts; a *comprehensive concept*. A <descrip> element that has a *type* attribute value of 'superordinateConceptPartitive' can also have a *target* attribute pointing to the entry identifier of the superordinate concept.

### **table**

The name of a table that is used to describe a concept. The table itself can be contained (or referred to as an external file) in a <refobject> in the back matter of a TBX document instance. The *target* attribute points to the <refobject> identifier. The element content shall be in plainText. NOTE If a direct pointer to the file that contains the table is preferred within the <termEntry>, rather than pointing to the back matter, use an <xref> instead.

### **temporallyRelatedConcept**

A concept related to the subject concept through a sequential relation involving events in time. A <descrip> element that has a *type* attribute value of 'temporallyRelatedConcept' can also have a *target* attribute pointing to the entry identifier of the related concept.

### **thesaurusDescriptor**

The term in a thesaurus that can be used to represent a concept in a terminological entry. A <descrip> element that has a *type* attribute value of 'thesaurusDescriptor' can also have a *target* attribute pointing to the ID of the <refObject> in the back matter of the TBX document instance that describes the thesaurus. The element content shall be in plainText. NOTE If a direct pointer to the thesaurus is preferred within the <termEntry>, rather than pointing to the back matter, use an <xref> instead.

### **unit**

A precisely-specified quantity in terms of which the magnitudes of other quantities can be stated, such as meter or volt. The element content shall be in plainText.

### **video**

A name or other description of recorded visual images used to represent or illustrate terminological information. A <descrip> element that has a *type* attribute value of 'video' can also have a *target* attribute pointing to the ID of a <refObject> in the back matter of the TBX document instance that describes and points to the video file. The element content shall be in plainText. NOTE If a direct pointer to the video file is preferred within the <termEntry>, rather than pointing to the back matter, use an <xref> instead.

## **descripNote**

### **contextType**

The characterization of a context according to a set of theoretical or pragmatic types. The content of the <descripNote> element when it has a *type* attribute value of 'contextType' shall be one of the following:

- definingContext
- explanatoryContext
- associativeContext
- linguisticContext
- metalinguisticContext
- translatedContext

### **definitionType**

The characterization of a definition according to a set of theoretical or pragmatic types. The content of the <descripNote> element when it has a *type* attribute value of 'definitionType' shall be one of the following:

- intensionalDefinition
- extensionalDefinition
- partitiveDefinition
- translatedDefinition

## **hi**

### **entailedTerm**

A term, found in the content of a text element such as <descrip type="definition">, that is defined in another terminological entry in the same terminological collection.

### **hotkey**

A key or a combination of keys on a computer keyboard that, when pressed at one time, performs a task (such as starting an application) as an alternative to using a mouse or other input device. Hotkeys are supported by many operating systems and applications (the hotkey is indicated by an ampersand in the term and by an underscore in the visual display). This type of <hi> element allows a hotkey to be identified as such in a text field such as a context

sentence. To identify a hotkey for the head term of an entry, use the `lionHotkey` value of the `type` attribute on a `<termNote>`.

#### **italics**

Indicates that the content enclosed in the `<hi>` element is rendered in italics.

#### **bold**

Indicates that the content enclosed in the `<hi>` element is rendered in boldface.

#### **superscript**

Indicates that the content enclosed in the `<hi>` element is rendered in superscript.

#### **subscript**

Indicates that the content enclosed in the `<hi>` element is rendered in subscript.

#### **math**

Indicates that the content enclosed in the `<hi>` element is mathematical in nature.

#### **ref**

##### **crossReference**

A pointer to another related location, such as another entry or another term. The `<ref>` element with a `type` attribute value of 'crossReference' shall also have a `target` attribute to point to the identifier of the entry or term.

##### **see**

An element that points to a location where additional information can be found. The `<ref>` element with a `type` attribute value of 'see' shall also have a `target` attribute to point to the location of the information.

#### **termCompList**

##### **hyphenation**

The `<termComp>` elements contain the parts of a term as it would be hyphenated, such as at the end of a line, according to a given set of rules.

##### **lemma**

The `<termComp>` elements contain the base form of the term or of the individual words that it comprises. A base form of a word or term, sometimes called a lemma, is the form that is used as the formal dictionary entry for the term. For nouns, the base form is frequently the nominative singular form (in languages that show variation by case). For adjectives it is the positive form of the adjective, as opposed to the comparative, and in some languages it is uninflected, whereas in others it is the masculine singular. For verbs, it is generally the infinitive.

##### **morphologicalElement**

The `<termComp>` elements contain the individual morphemes that comprise the term.

##### **syllabification**

The `<termComp>` elements contain the individual syllables that comprise the term.

##### **termElement**

The `<termComp>` elements contain any other logically significant portion of the term.

#### **termNote**

##### **abbreviatedFormFor**

A link used to identify a relation between a term that is an abbreviated form, and its full or expanded form. It may be coupled with a `target` attribute to point to the entry of the full or expanded form. For example, the term WWW could have the following tag to point to its full form: `<termNote type="abbreviatedFormFor" target="tid-worldwideweb">World Wide Web</termNote>`. The actual content of the element would be the text *World Wide Web*, which allows the user to see what term is the full form of the current term without having to go to the location of the full form in the file (the location would normally be somewhere within the current entry, since the abbreviated form

and the full form represent the same concept). The element content shall be expressed in `basicText`. The identification of a relation between an abbreviated term and its full form indicates which term in a terminological entry is the full form of the abbreviated term. In many languages, English abbreviations are commonly used, but the English full form may be avoided in favor of an equivalent full form in the language in question. In this case, the English abbreviation and the target language full form will both appear in the target language section of the entry.

#### **administrativeStatus**

The status of a term within a certain working environment. The data-category `administrativeStatus` should be linked, through a *target* attribute, to a `<refObject>` in the back matter that identifies the administrative organization that determined the status. The element `<termNote>` when its *type* attribute has the value 'administrativeStatus' shall contain one of the following values:

- `admittedTerm-admn-sts`
- `deprecatedTerm-admn-sts`
- `legalTerm-admn-sts`
- `preferredTerm-admn-sts`
- `regulatedTerm-admn-sts`
- `standardizedTerm-admn-sts`
- `supersededTerm-admn-sts`

NOTE The status of terms with respect to a standardization process is covered by `normativeAuthorization`.

#### **animacy**

The characteristic of a word indicating that in a given discourse community, its referent is considered to be alive or to possess a quality of volition or consciousness. The value of the `<termNote>` element when it has a *type* attribute value of animacy shall be one of: `animate`, `inanimate`, `otherAnimacy`.

#### **antonymTerm**

A term that designates a concept that is opposite in meaning to the concept represented by the entry containing the `<termNote type='antonymTerm'>` element. A `<termNote>` element that has a *type* attribute value of 'antonymTerm' can also have a *target* attribute pointing to the term ID of the term that represents the antonym concept. The display term in the content of this element shall be expressed in `basicText`.

#### **directionality**

A data-category that, through the value of its *target* attribute, indicates the preferred translation of the term in the `<term>` element of the current `<tig>` from among the available translations in the entry. The *target* attribute should point to the identifier of the `<tig>` corresponding to the target term. In the case of the values `monodirectional` and `bidirectional`, the target term is the preferred term. The content of the `<termNote>` element when it has a *type* attribute value of 'directionality' indicates the scope of directionality, and shall be one of the following:

- `monodirectional`: The translation is only preferred going from the source language (current `<tig>`) to the target language (target `<tig>`).
- `bidirectional`: The translation is preferred in both directions, i.e. the term in the target `<tig>` is a preferred translation of the term in the current `<tig>` and vice-versa.
- `incommensurate`: The relation between the term in the current `<tig>` and the term in the target `<tig>` as translations is imperfect and should be used with care. In this case, the concepts are not equal. A `<termNote type='transferComment'>` should be used to provide information about the difference between the concepts.
- `undetermined`: The scope of directionality has not been determined.

#### **etymology**

Information on the origin of a word and the development of its meaning. The element content shall be expressed in `noteText`.

#### **falseFriend**

A term in another language that appears to be semantically or morphologically similar to the subject term in the entry, but that does not represent the same concept. For example, in French, the term *réaliser* means "to accomplish", and therefore, it is a false friend of the English term *realize* (i.e. to become conscious of). A `<termNote>` element with the *type* attribute value of 'falseFriend' can also have a *target* attribute whose value is a pointer to a term in another entry. The element content shall be expressed in `basicText`.

**frequency**

The relative commonness with which a term occurs. The content of the <termNote> element when it has a *type* attribute value of 'frequency' shall be one of: commonlyUsed, infrequentlyUsed, rarelyUsed. The assessment of frequency of a term can be based on subjective criteria, or it can reflect computer analysis of text corpora.

**geographicalUsage**

A code identifying any specific geographical area where the term is used, for instance, to differentiate between UK English and US English. Preferred values are language codes combined with geographical codes, from IETF RFC 4646 or its successor, as identified in IETF BCP 47. For example: en-US for American English. The element content shall be expressed in plainText.

**grammaticalGender**

The gender of the word. The value of the <termNote> element when it has a *type* attribute value of 'grammaticalGender' shall be one of: masculine, feminine, neuter, or otherGender.

**grammaticalNumber**

A property of a term that indicates the number of objects referred to by the term. The content of the <termNote> element when it has a *type* attribute value of 'grammaticalNumber' shall be one of: singular, plural, dual, mass, or otherNumber. For example, the terms rice and bread are mass nouns because there are no separate singular and plural forms. NOTE If the singular and the plural forms of a term do not designate the same concept, they should be recorded in separate entries.

**grammaticalValency**

The range of syntactic elements either required or permitted by a lexical unit. The valency of eat includes a subject ("I" in "I am eating") and an object ("cheese" in "I am eating cheese"). It can thus be said to be "bivalent". The element content shall be expressed in plainText.

**homograph**

A word that is spelled like the subject term in the entry, but that has a different meaning, and origin. Homographs are terms that have the same written form, but are derived from different etymological origins, for example, a port for boats and a port for computer peripherals. A <termNote> element having a *type* attribute value of 'homograph' can also have a *target* attribute to point to the <tig> of the entry that describes the homograph. The element content shall be expressed in basicText.

**language-planningQualifier**

A qualifier assigned to a provisional term within a language planning or descriptive terminology environment. The content of the <termNote> element when it has a *type* attribute value of 'language-planningQualifier' shall be one of: recommendedTerm, nonstandardizedTerm, proposedTerm, or newTerm.

**lionHotkey**

(Localization hot key) The letter in a word that acts as a hotkey in a computer application. A hotkey is a key or a combination of keys on a computer keyboard that, when pressed at one time, performs a task (such as starting an application) as an alternative to using a mouse or other input device. Hotkeys are supported by many operating systems and applications. The hotkey is indicated by an ampersand in the term in software interface files, and by an underscore in the actual visual display. The element content shall be expressed in plainText.

**normativeAuthorization**

A term status qualifier assigned by an authoritative body, such as a standards body or a governmental entity with a regulatory function. The content of the <termNote> element when it has a *type* attribute value of 'normativeAuthorization' shall be one of:

- standardizedTerm
- preferredTerm
- admittedTerm
- deprecatedTerm
- supersededTerm
- legalTerm
- regulatedTerm

The name of the authoritative body should be documented in the terminology repository, and referenced in the TBX document instance by using a <refObject>.

### partOfSpeech

A category assigned to a word based on its grammatical and semantic properties. For example: noun, verb, or adjective. The element content shall be expressed in plainText, and preferably be limited to a set of picklist values. Only one value shall be permitted for each <term> in a given entry. For example, the following values are permitted in the *TBX-Basic* TML:

- noun
- verb
- adjective
- adverb
- properNoun
- other

### processStatus

The status of a term with respect to its advancement within the terminology management process. The content of the <termNote> element when it has a *type* attribute value of 'processStatus' shall be one of: unprocessed, provisionallyProcessed, or finalized.

### pronunciation

The representation of the manner by which a term is spoken. In TBX, pronunciation shall be represented in the International Phonetic Alphabet. A given term can have more than one pronunciation depending on the geographical area in which the term is used. In this case use a separate <tig> or <ntig> element for each pronunciation, and indicate the geographical area for each by using a <termNote> element with a *type* attribute value of 'geographicalUsage'. The element content shall be expressed in basicText.

### proprietaryRestriction

A restriction placed on a term for the purpose of protecting the right of an organization to the exclusive use of the term. The content of the <termNote> element when it has a *type* attribute value of 'proprietaryRestriction' shall be one of: trademark, serviceMark or tradeName. NOTE The content value "copyright" is not included because copyright only applies to larger text units, such as books, not to terms.

### register

Classification indicating the level of language assigned to a term. The content of the <termNote> element when it has a *type* attribute value of 'register' shall be one of:

- colloquialRegister
- neutralRegister
- technicalRegister
- in-houseRegister
- bench-levelRegister
- slangRegister
- vulgarRegister

### shortFormFor

A linking element used to identify a relation between the subject term when it is a short form, and its full or expanded form. It may be coupled with a *target* attribute to point to the full or expanded form. Normally both the short form and the full form will be contained in the same entry, but each in its own <tig>, because they represent the same concept. For example, the term United States could have the following markup to point to the full form term United States of America: <termNote type="shortFormFor" target="tid-unitedstatesofamerica">United States of America</termNote>. The content of the element would be the target term, expressed in basicText, which allows the user to see the target term without going to its location in the entry. The identification of a relation between a short form and its full form indicates which term in a terminological entry is the full form of the short form.

### temporalQualifier

A property of a term with respect to its use over time. The content of the <termNote> element when it has a *type* attribute value of 'temporalQualifier' shall be one of: archaicTerm, outdatedTerm, obsoleteTerm.

**termLocation**

A location in a document, computer file, or other information medium, where the term frequently occurs, such as a user interface object (in software), a packaging element, a component in an industrial process, and so forth. The element content shall be expressed in plainText, and preferably be restricted to a set of picklist values. The following picklist values are recommended for software user interface locations in a Windows environment.

- checkBox
- comboBox
- comboBoxElement
- dialogBox
- groupBox
- informativeMessage
- interactiveMessage
- menuItem
- progressBar
- pushButton
- radioButton
- slider
- spinBox
- tab
- tableText
- textBox
- toolTip
- user-definedType

**termProvenance**

Classification of a term according to the methodology employed in creating the term. The content of the <termNote> element when it has a *type* attribute value of 'termProvenance' shall be one of: transdisciplinaryBorrowing, translingualBorrowing, loanTranslation, or neologism.

**termStructure**

The structure of the components of a term, indicated using square brackets. For example, e.g. : [bank statement] [total] versus [bank] [statement total]. The element content shall be expressed in plainText.

**termType**

A categorization of a term with respect to formal representations of terms or where or how they are used, for example, acronym. The content of the <termNote> element when it has a *type* attribute value of 'termType' shall be one of the following:

- abbreviation
- acronym
- clippedTerm
- commonName
- entryTerm
- equation
- formula
- fullForm
- initialism
- internationalism
- internationalScientificTerm
- logicalExpression
- partNumber
- phraseologicalUnit
- transcribedForm
- transliteratedForm
- shortForm
- shortcut
- sku
- standardText
- string
- symbol
- synonym
- synonymousPhrase
- variant

Refer to ISO 12620 for definitions of the above term types. The value shortcut refers to a string that acts as a keyboard shortcut on computers, for example, *Ctrl + V* is a term that represents a keyboard shortcut for the concept of *paste*.

**timeRestriction**

The indication of a period of time during or since which a term was or has been subject to specified usage. For example, several European countries have redefined the requirements for certain university degrees in recent years. If, for instance, the requirements for a baccalaureate degree changed from three to four years in 1993, then any terminology entry defining the term used to designate this degree would have to specify the time restriction affecting the definition, which would be different before 1993 and after 1993. The element content shall be expressed in plainText.

**transferComment**

Note included in a term entry providing information on the degree of equivalence, directionality or other special features affecting equivalence between a term in one language and another term in a second language. This element can be used in conjunction with the <termNote type="directionality"> element to indicate the preferred translation of the term in question. The <termNote> element with a *type* attribute value of 'transferComment' can also have a *target* attribute to point to the term that is the object of the transfer comment.

**usageNote**

A note containing information on the usage of the term.

**transac**

**transactionType**

A record that indicates the stage of the entry within the overall process of creation, approval, and use of a terminology entry.

**transacNote**

**usageCount**

The number of times an entry has been accessed.

**responsibility**

The name of a person or entity associated with a terminology management transaction. The <transacNote> element with a *type* attribute value of 'responsibility' can also have a *target* attribute to point to the identifier of a bibliographic entry in a <refObject> that contains information about the person or entity.

**xref**

**corpusTrace**

A notation indicating the location of an external contextual resource.

**externalCrossReference**

A term or other terminological resource that is external to the terminological database or TBX document instance.

**xAudio**

External audio file.

**xGraphic**

External graphic file.

**xVideo**

External video file.

**xSource**

An external source of information such as the source of a definition or context sentence.

**xMathML**

A mathematical concept expressed in MathML.