**ISO/TC 37/SC 3**

Date: 2001-12-5

**ISO/DIS 16642**

ISO/TC 37/SC 3/WG

Secretariat: DIN

# Computer applications in terminology - Terminological markup framework (TMF)

***Terminotique - Plate-forme pour le balisage de terminologies informatisées***



---

**Content**

**Foreword**

**Introduction**

**1 Scope**

**2 Normative references**

**3 Terms and definitions**

**4 General principles**

**5 A generic model for describing linguistic data and its application to terminology**

5.1 Introduction

5.1.1 General principles

5.1.2 Examples

5.2 Generic representation of structural levels and information units

5.3 The terminological meta-model

5.4 Designing representations of terminological data on the basis of the meta-model

5.5 Interchange, dissemination and interoperability

5.6 XML canonical representation of the generic model

5.6.1 Introduction

5.6.2 Example

5.6.3 A thorough description of the GMT format

5.6.3.a) The tmf element

5.6.3.b) The struct element

5.6.3.c) The feat element

5.6.3.d) The brack element

5.6.3.e) The annot element

5.7 Representing languages in a terminological data collection

6 Defining a terminological markup language (TML)

6.1 General

6.2 Defining interoperability conditions

6.3 Implementing a TML

6.3.1 Introduction

6.3.2 Expanding the structural skeleton (XML outline of a TML)

6.3.3 Anchoring data categories on the TML XML outline

6.3.3.a) Styles and vocabulary

6.3.3.b) Constraints on datatypes for information units

6.3.3.c) External markup modules

6.3.4 Implementing annotations

6.3.5 Implementing brackets

6.3.6 Namespaces

Annex A (normative): XML schema of the GMT format

Annex B (normative): The MSC Terminology Markup Language

B.1 Introduction

B.2 An example of an MSC XML document

B.3 Expansion trees

B.4 Data categories

B.5 Conclusion

Annex C (normative): The Geneter Terminology Markup Language

C.1 Introduction

C.2 Example : specification of a Geneter subset as a TML

C.2.1 Geneter subset

C.2.2 XML encoding of an entry

C.2.3 Geneter vocabulary and style

C.3 Global Information Section (GIS)

C.4 Terminological entry (TE)

C.4.1 Data category types

C.4.2 Characteristics of a Geneter data category: name, attributes, content model

C.4.3 General attributes

C.4.4 Container

C.4.5 Content models

C.4.5.1 mixed content

C.4.5.2 composite content

C.4.6 Pointers

C.4.7 Structure of a terminological entry

C.4.8 Example of terminological entry

C.4.9 The tree structure of a terminological entry

C.4.9.1. Geneter synopsis

C.4.9.2 Non 12620 data categories

C.4.9.3 Entities for content models

C.4.9.4 Entities for suggested picklists

C.5 Complementary Information (CI)

C.5.1 Geneter Complementary Information types

C.5.2 A mechanism for extending Complementary Information

C.6 Geneter restriction and extension

C.6.1 Rules to constitute a subset :

C.6.2 Different types of subset :

C.6.3 Blind subset

C.6.4 Building a subset : an example

C.6.4.1 A data sample

C.6.4.2 1st step : Analysis of the elements

C.6.4.3 2nd step : The projection scheme from the source format to Geneter

C.6.5 Geneter extensions and negotiated interchange

Annex D (Informative): Conformance of terminological data to TMF

D.1 Sample terminological data

D.2 Description of content of elements

D.3 Conformance to TMF

D.3.1 Meta-model specification

D.3.2 Data Category Specification (DCS)

D.3.3 Content mappings

D.3.4 TMF-conformant XML-representation (GMT)

---

**Foreword**

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

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

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 International Standard may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

International Standard ISO 16642 was prepared by Technical Committee ISO/TC 37, Terminology and other language resources, Subcommittee SC 3, Computer applications in terminology.

Annexes A, B and C form an integral part of this International Standard. Annex D is for information only.

**Introduction**

Terminological data are collected, managed and stored in a wide variety of systems, typically in applications, i.e., various kinds of database management systems, ranging from personal computer applications for individual users to mainframe term-bank systems operated by major companies and governmental agencies. Termbases are comprised of various sets of data categories and based on various kinds of data models. Terminological data often need to be shared and reused in a number of applications, and this sharing is usually accomplished using intermediate formats. To facilitate co-operation and to prevent duplicate work, it is important to develop standards and guidelines for creating and using terminological data collections as well as for sharing and exchanging data.

The meta-model fits into an integrated approach to be used in analysing existing terminological data collections and in designing new ones, which are typically processed using relational or text-based data management systems. Terminological data collections **can** also be stored as structured documents with markup based on formats that are

typically defined using Standard Generalized Markup Language (SGML, ISO 8879:1986) or eXtensible Markup Language (XML), which is based on SGML, but amended for use on the World Wide Web (WWW) by the World Wide Web Consortium (W3C). An integrated approach eases the task of importing data from a flat file with markup into a database and the task of exporting data from a database to a structured document. Yet another motivation for an integrated approach, as opposed to entirely separate approaches for databases and structured documents, is that XML-based formats are now being processed in new ways, similar to traditional database management systems. For example, XML files are being queried and updated directly without importing data into traditional database environments.

This integrated approach to analysis and design consists of two levels of abstraction. The first (and most abstract) level of the integrated approach is the meta-model level. The meta-model level, which could also be called the abstract conceptual data model level, supports analysis and design at a very general level. The second level is the data model level. At this level, the designer of the terminological data collection has the possibility to make various choices, based on real-life needs. First, designers must determine the form of representation most appropriate for their terminological data, addressing the following choices:

- Whether to use a relational database or a flat file with markup;
- Whether the terminological data will be used primarily for queries and updates and be represented in some database management system;
- What system to use if this is the case;
- Whether the data will be used primarily for sharing and interchange and be represented in a flat file with markup.

For the application of this International Standard it is assumed that all flat files will use XML markup.

Once the choice between a database management system and XML has been made, a data model must be chosen. For a relational database, a typical method of describing a data model is an entity-relationship diagram. For an XML format, a typical method of describing a data model is a Document Type Definition (DTD). An alternative method, using what is called an "XML schema", is provided by the W3C. In the future, other, it **may** be possible that more abstract methods of describing an XML format will be used.

A specific implementation of the meta-model for terminology markup expressed in XML is called a terminological markup language (or TML), which **can** be described on the basis of a limited number of characteristics, namely:

- How the TML expresses the structural organisation of the meta-model (i.e. the *structural skeleton* of the TML);
- The specific data categories used by the language and how they relate to the meta-model;
- The way in which these categories **can** be expressed in XML and thus anchored on the structural skeleton of the TML, that is the XML *style* of any given data category;
- The *vocabularies* used by the TML to actually express those various informational objects as XML elements and attributes according to the corresponding styles.

Some of the examples in the present document are instances of the MSC and Geneter formats as described in annex B and annex C of this document.

Some further documents, comprising in particular transformation experiments between Geneter and MSC by way of the GMT pivot format can be found at http://www.loria.fr/projets/TMF.

**1 Scope**

Effective use and re-use of a variety of multilingual terminological resources is facilitated by a single high-level data model that supports analysis and design of both databases and intermediate formats. This International Standard specifies a model that has been designed for the purpose of providing guidance on the basic principles for representing terminological data, as well as for describing specific terminological markup languages.

This International Standard is designed to support the development and use of computer applications for terminological data and the exchange of such data between different applications. Standardisation of data categories and methods for the definition of data structures are specified in ISO 12620 and other related International Standards.

This International Standard specifies a framework designed to provide guidance on the basic principles for representing data recorded in terminological data collections. This framework includes a meta-model and methods for describing specific terminological markup languages (TMLs) expressed in XML. The mechanisms for implementing constraints in a TML are defined in this International Standard, but not the specific constraints for individual TMLs (which **can** be the subject of further standardizations), exept for the three TMLs defined in the annexes

This International Standard also defines the conditions that allow the data expressed in one TML to be mapped onto another TML and specifies a generic mapping tool, GMT, for this purpose.

In addition, this International Standard also describes a generic model for describing linguistic data.

## 2 Normative references

The following normative documents contain provisions, which, through reference in this text, constitute provisions of this International Standard. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this International Standard are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards.

ISO/IEC 639, *Information technology - ISO* 639:1988, Code for the representation of names of languages.

ISO 639-2:1998, Code for the representation of names and languages-part 2:Alpha-3 code.

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

ISO 690:1987, *Documentation - Bibliographic references - Content, form and structure*.

ISO 690-2:1997, *Information and documentation - Bibliographic references - Part 2: Electronic documents or parts thereof*

ISO 704:1999, *Principles and methods of terminology.*

ISO 1087-1:2000, *Terminology work - Vocabulary - Part 1: Theory and application*

ISO 1087-2:2000, *Terminology work - Vocabulary - Part 2: Computer applications.*

ISO 2788:1986, *Guidelines for the establishment and development of monolingual thesauri*

ISO 3166-1:1997, *Code for the representation of names of countries and their subdivisions - Part 1: Country codes*

ISO 5964:1985, *Guidelines for the establishment and development of multilingual thesauri*

ISO 8601:1988, *Data elements and interchange formats - Information interchange - Representation of dates and times.*

ISO 8879:1986 (*SGML*) as extended by TC2 (ISO/IEC JTC 1/SC 34 N 029:1998-12-06) to allow for *XML*.

ISO/IEC 10646-1:1993, *Information technology-Universal Multiple-Octet Coded Character Set (UCS)-Part 1: Architecture and basic multilingual plane.*

ISO 12083:1994, *Information and documentation - Electronic manuscript preparation and markup*.

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

ISO 12620, *Computer applications in terminology - Data categories*.

## 3 Terms and definitions

For the purposes of this International Standard, the terms and definitions given in ISO 1087-1, ISO 1087-2 and the following apply.

**3.1**
**complementary information (CI)**
information supplementary to that described in terminological entries and put in factor across the terminological data collection
**Note:** typical content of CI are: domain hierarchies, institution descriptions, bibliographical references,etc.

**3.2**
**data category**
result of the specification of a given data field [ISO 1087-2:2000], (i.e. a type of data field, such as *definition*)
**Note:** ISO 12620 is an inventory of data categories, i.e., a data category registry or DCR

**3.3**
**DCR; data category registry**

data category specification used as a normative reference for the description of a TML
**Note:** ISO 12620 is a typical DCR in the context of the present standard

**3.4**
**DCS; data category specification**
component of a TML's definition that constrains its informational content, e.g., by specifying which data categories are allowed and how each data category **can** be used

**3.5**
**expansion tree**
list of XML elements together with their organization that implement a level of the meta-model in a given TML

**3.6**
**global information section (GIS)**
technical and administrative information applying to the entire data collection
**Example:** title of the data collection, revision history

**3.7**
**information unit**
elementary piece of information attached to a level of the meta-model

**3.8**
**language section (LS)**
part of a terminological entry containing information related to one language
**Note:** One terminological entry **may** contain information on one, two or more languages

**3.9**
**object language**
language of the terminological information under description at a given place of a terminological entry

**3.10**
**structural level**
level of the meta-model to which one or more information units **can** be attached

**3.11**
**structural skeleton**
set of XML elements which, in a given TML, results from the expansion of the meta-model

**3.12**
**style**
properties associated to a data category that determine how it **may** be expressed in XML

**3.13**
**term component section (TCS)**
section of a term section giving information about components of a term

**3.14**
**term section (TS)**
part of a language section giving information about a term
**Example:** usage of a term, term elements

**3.15**
**terminological data collection (TDC)**
collection of data containing information on concepts of specific subject fields

**3.16**
**terminological entry (TE)**
entry containing information on terminological units (i.e., subject-specific concepts, terms, etc.)
**Note:** Every element in the terminological entry **can** be linked to complementary information, to other entries and to other elements in the same entry

**3.18**
**TML; terminological markup language**
XML application for describing terminological data collection compliant with the constraints expressed in this standard

**3.19**

**vocabulary**
set of strings used to implement a data category according to a style

**3.20**
**working language**
language used to describe any textual content in a terminology data collection

**3.21**
**XML; extensible markup language**
universal format for structured documents and data on the WWW; a particular subset of SGML

**3.22**
**XSLT; extensible stylesheet language transformation**
language for transforming XML documents into other XML documents

## 4 General principles

Describing a specific terminological markup language (TML) **can** be seen as a process involving several knowledge sources which interact with one another at various levels. This process allows to specify two important aspects of a TML:

- The informational properties of the language, that is its capacity to represent a given piece of information related to terminological description;
- The way the language **can** be expressed, for instance as an XML document.

Figure 1 represents the various knowledge sources that are part of the definition of TMF and that **can** lead to the full definition of a given terminological markup language (TML).

Two of those knowledge sources are shared by all TMLs and **can** be seen as reference material for this standard:

- The **meta-model** describes the basic hierarchy of information levels to which any TML **shall** conform as defined in this standard;
- A **Data Category Registry** (DCR) is a set of data category specifications on which any specific TML **shall** rely for creating its own data category set. For the application of this International Standard, ISO 12620 forms a reference DCR for any information unit to be used in the definition of a TML.

Two other knowledge sources are used to define the specific information units of a given TML from the point view of both its informational properties and its representation in XML:

- The **Data Category Specification** (DCS) describes the set of data categories that **can** be used within a given TML. The DCS **can** comprise both a subset of the $DCR_{ref}$'s and the addition of any idiosyncratic data categories needed for a specific application;
- The **Dialectal Specification** (Dialect) includes the various elements needed to describe a given TML as an XML document. These elements comprise expansion trees and data category instantiation styles, together with their corresponding vocabularies.

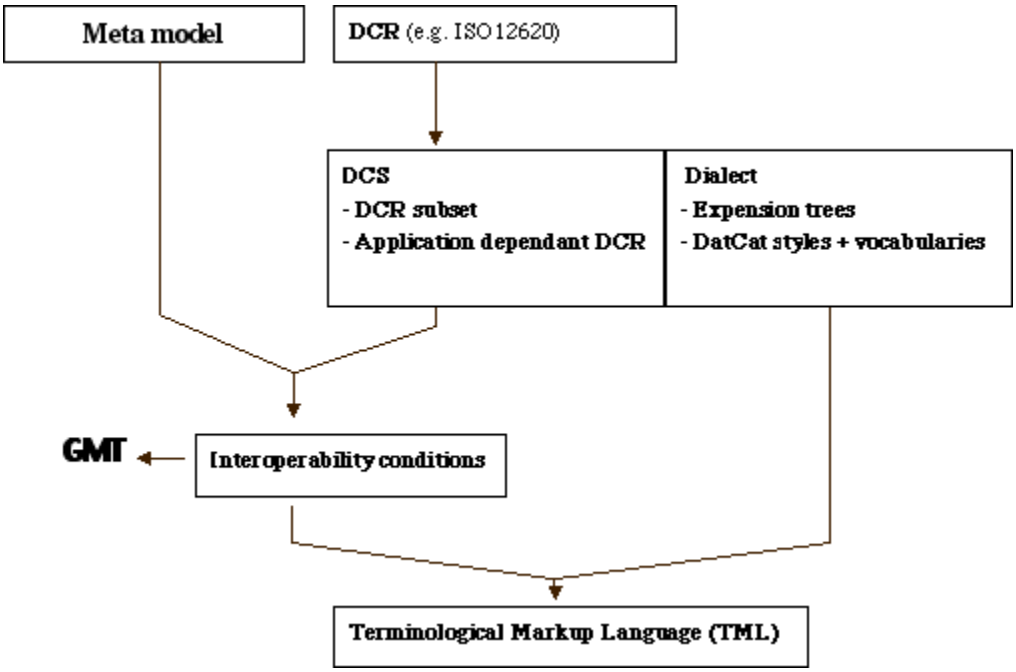



**Figure 1: The various knowledge sources involved in the description of a TML**

**Interoperability principle**

The combination of the meta-model and a given DCS is enough to define conditions of interoperability, encompassing its full informational properties from a terminological point of view. Any information structure that corresponds to such conditions has a canonical expression as an XML document using the GMT (Generic Mapping Tool) representation. The interoperability between two different TMLs depends solely on their compatibility at that level (cf. figure 2).

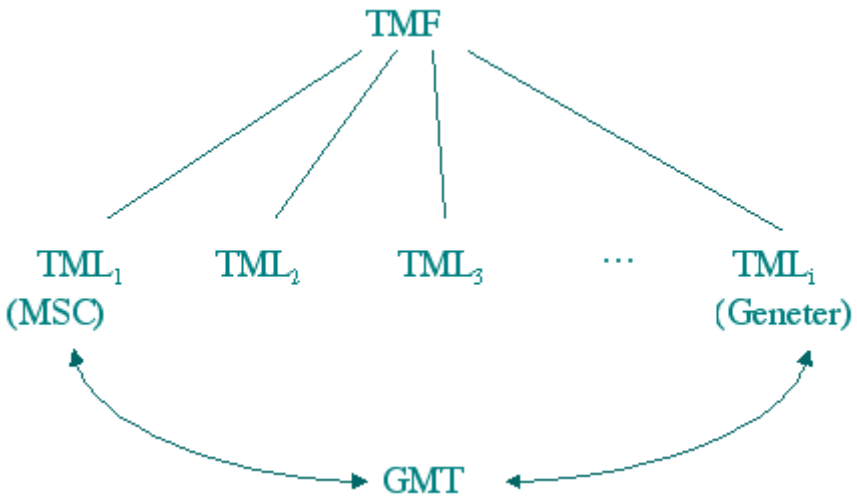



**Figure 2: Interoperability between two TMLs using GMT.**

More precisely, the interoperability between two TMLs **can** be defined by comparing their specifications. Under the condition that two TMLs are based upon the same DCS, any terminological data collection expressed in either TML **can** be transformed, without any loss of information, into a collection represented using the other format. In particular, the additional specifications that are needed to express interoperability conditions in the form of an XML implementation of a TML do not influence the level of interoperability between two languages. The preceding interoperability principle leads to two immediate consequences:

- It justifies the role of GMT as a pivot language when transforming any data expressed in a TML $TML_i$ into data expressed in a TML $TML_j$;
- When two TMLs are based upon two different DCSs, it provides a framework for identifying what information **can** be transformed between one format and another and what will be lost during the transformation (weak interoperability condition).

The combination of interoperability conditions with dialectal information provides the information necessary to express a

given TML in XML.

## 5 A generic model for describing linguistic data and its application to terminology

### 5.1 Introduction

#### 5.1.1 General principles

Linguistic structures exist in a wide variety of formats ranging from highly organized data (e.g. terminological data bases) to loosely structured information (for example, text lightly annotated for names, dates etc.). The representation of annotated corpora is not hierarchical, but based on the expression of multiple views representing various levels of linguistic information, usually pointing to primary data (e.g. part of speech tagging) and sometimes to one another (e.g. reference annotation based on basic phrase structure annotation).

The following model identifies a class of document structures which could be used to cover a wide range of linguistic annotation formats, and provides a framework which can be applied using XML.

Each type of document structure is described by means of a three-tiered information structure that describes:

- a *meta-model*, which represents a hierarchy of structural nodes which are relevant for linguistic description;
- specific *information units*, which can be associated with each structural node of the meta-model;
- relevant *annotations*, which **can** be used to qualify some parts of the value associated with a given information unit.

Each structural node **can** be qualified by a group of basic or compound information units. A basic information unit describes a property that **can** be directly expressed by means of a data category. A compound information unit corresponds to the grouping at one level of several basic information units, which, taken together, express a coherent unit of information.

For instance, a compound information unit **can** be used to represent the fact that a transaction **can** be a combination of a transaction type, a responsibility, and the transaction date.

Basic information units, whether they are directly attached to a structural node in the structural skeleton or within a compound information unit, **can** take two non-exclusive types of values:

- An atomic value corresponding either to a simple type (in the sense of XML Schemas) such as a number, string, element of a picklist, etc., or to a mixed content type in the case of annotated text;
- A reference to a structural node within the meta-model in order to express a relation between it and the current structural node.

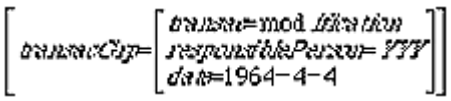
Basic and compound information units **can** be abstractly represented as information unit structures associated with specific structural nodes in the structural skeleton. For instance, a Geneter sub-document identified as <Owner>UHB</Owner> **can** be modelled as a basic information unit in the following way:

```
[owner = UHB]
```

Similarly, the following MSC sub-document

```
<transacGrp>
     <transac>modification</transac>
     <transacNote type="responsiblePerson">YYY</transacNote>
     <date>1964-4-4</date>
</transacGrp>
```

can be modelled as:



The preceeding model for information units is to be completed by a last level of information representation, which corresponds to the association of semantic information to subparts of information unit values. Such *annotations* typically occur when one wants to identify, within a terminological definition, specific references to information involving genus and/or differentia. For instance in a definition for *lead pencil*:

```
<definition>
     <broaderConcept>pencil</broaderConcept> whose
     <characteristic>casing</characteristic> is fixed around a central
     <characteristic>graphite </characteristic> medium which is <characteristic>used
     for writing or making marks</characteristic>
</definition>
```

Such information, also known as *mixed content* in XML, **cannot** be directly represented as a feature structure and will be directly expressed in the following GMT representation ("annot" element).

#### 5.1.2 Examples

To illustrate how a TDC can be analysed as an abstract structure, let us consider a simple terminological entry expressed as an XML document conformant to MSC specifications:

```
<?xml version="1.0"?>

<martif type="MSC" lang="en">
     <text>
          <body>
               <termEntry id="ID67">
                    <descrip type="subjectField">manufacturing</descrip>
                    <descrip type="definition">A value between 0 and 1 used
                    in ...</descrip>
                    <langSet lang="en">
                         <tig>
                              <term>alpha smoothing factor</term>
                              <termNote type="termType"
                              datatype="picklistVal">fullForm</termNote>
                         </tig>
                    </langSet>
                    <langSet lang="hu">
                         <tig><term>Alfa ...</term></tig>
                    </langSet>
               </termEntry>
          </body>
     </text>
</martif>
```

The document represented above **can** be mapped to the abstract model described in this section by identifying a structural skeleton corresponding to the meta-model and by associating the corresponding information units with each structural node in the structural skeleton, as shown in figure 3.

Here, data categories can be mapped onto corresponding data categories specified in ISO 12620, that is:

| **Data category** | **ISO 12620 number** | **ISO 12620 name** |
|---|---|---|
| id | A10.15 | entry identifier |
| subjectField | A04 | subject field |
| definition | A05.01 | definition |
| lang | A10.07.01 | language identifier |
| Term | A01 | term |
| termType=fullForm | A02.01.07 | full form |

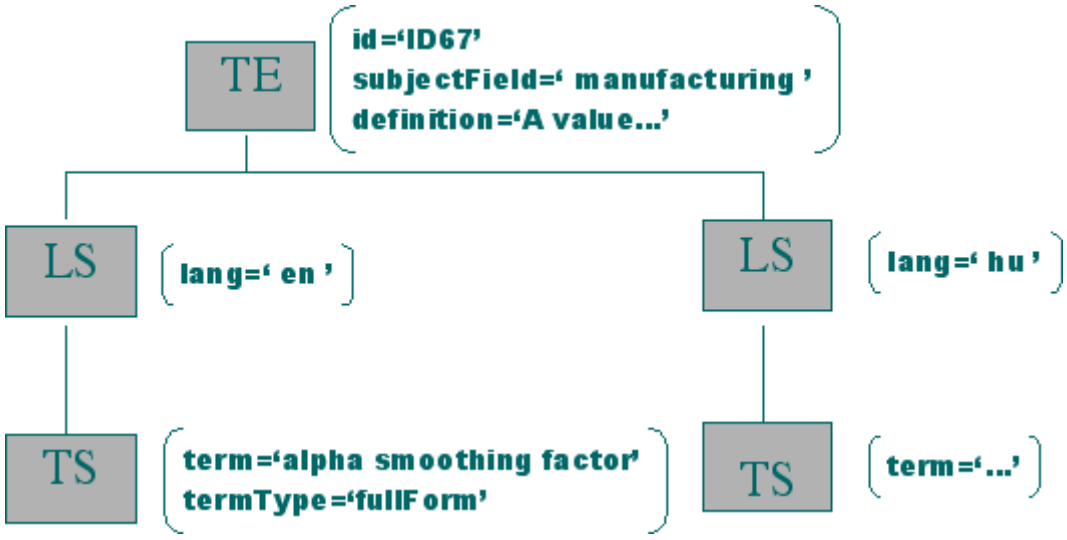



**Figure 3: Mapping a document to the abstract model**

#### 5.2 Generic representation of structural levels and information units

Linguistic data can be represented with a generic architecture that consists of a graph of elementary *structural nodes* to which one or more *information units* are attached. This architecture is shown in a UML diagram (figure 4). The diagram expresses the relationship between the following defined classes:

- *Structural Node*: a class containing one attribute (LevelName) that identifies objects of this type in the context of a given Linguistic Resource (LR) format (e.g. TE/Terminological Entry for the representation of TermBanks);
- *Information Unit*: a class containing three attributes to identify objects of this type in relation to a given Data Category (IUName; e.g. Definition, PartOfSpeech etc.), describe a type for its content (C_type) and provide the actual content value (C_value).

**Note:** the value of C_type can either belong to the set of simple types as defined in [XML Schemas part 2: Datatypes] or be MIXED (see below).

Objects of these two classes can be related in the following ways:

- Dependency: indicates that a structural node is related to another structural node by a hierarchical link. There is no constraint on the number of links and the structure of the network that those links create (tree, directed acyclic graph etc.);
- HasContent: relates a Structural Node to Information Units (e.g. a definition attached to a Terminological Entry). An instance of an Information Unit is attached to one and only one Structural Node;
- Refinement: relates an Information Unit (refined IU) to other Information Units (refining IU) that provide complementary information to the refined IU (e.g. a /note/ refining a /definition/). A refining IU is related to one and only one refined IU.

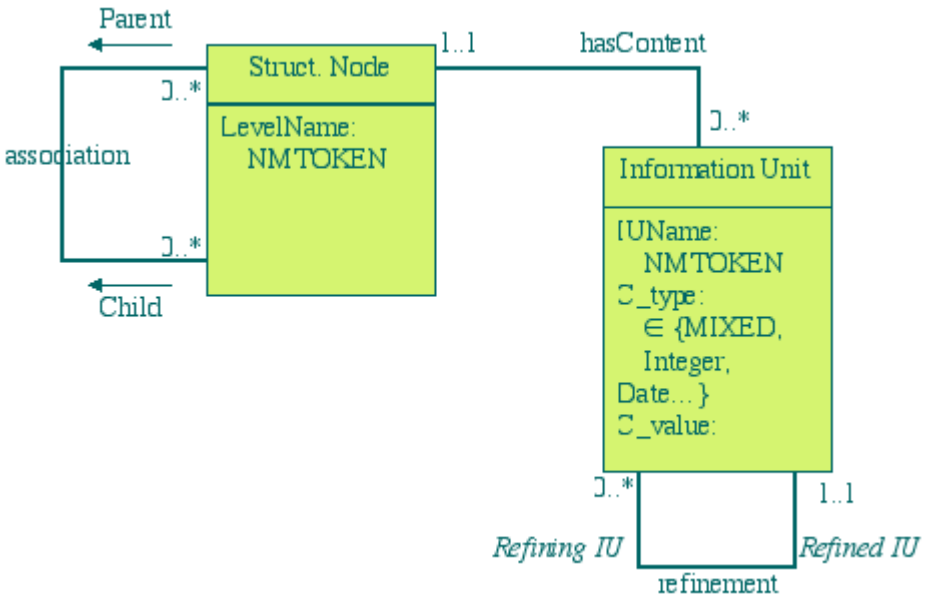


**Figure 4: A UML diagram for structural nodes and information units**

**MIXED type**

The MIXED type is an ordered combination of textual content and information units, corresponding to any kind of annotated content. It can be represented in UML by means of the agglomeration operator, as shown in figure 5.

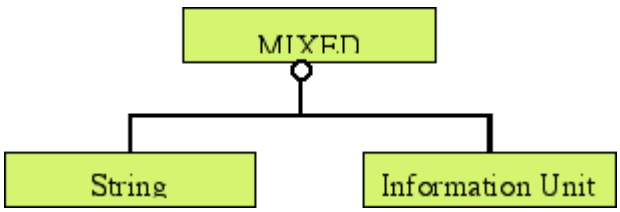



**Figure 5: The MIXED object class.**

Keeping this rather general definition leads to the fact that annotations in textual content can be refined by other information units (for instance to indicate when and by whom the annotation has been made).

#### 5.3 The terminological meta-model

The terminological meta-model is based on guidelines concerning the methods and principles of terminology management involving the production of terminological entries as described in ISO 704. One of the most important characteristics of a terminological entry - compared to a lexicographical entry - is its concept orientation. A terminological entry treats one concept in a given language and, in the case of multilingual terminological entries, one or more totally or partially equivalent concepts in (an)other language(s), whereas a lexicographical entry contains one lemma [the base form of a single lexical unit] and one or more definitions (representing different meanings) in one or more languages. The meta-model presented in this standard gives guidelines for terminological entries. It is understood in this context that although lexicographical entries typically contain word or lexeme related information (such as part of speech, gender, etc.), some of this lexicographical information is frequently also included in terminological entries.

A terminological data collection comprises global information about the collection and a number of entries. Each entry performs three functions:

- It describes one concept or two or more totally or partially equivalent concepts in one or more languages;
- It lists the terms that designate the concepts;
- It describes the terms themselves.

Each entry **can** have multiple language sections, and each language section **can** have multiple terminological units. Each data element in an entry **can** be associated with various kinds of descriptive and administrative information. In addition, there are various other resources that are not part of any one entry, but that **can** be linked to one or more entries. Such resources include bibliographic references, descriptions of ontologies, and binary data such as images that illustrate concepts.

By instantiating the generic architecture described previously, the terminological meta-model is extended by describing seven objects from the Structural Node class (with a number of dependency relations realised through use of the is_a relation):

- **TDC (Terminological Data Collection):**
  Top level *container* for all information contained in a terminology system. Generally used as a *container* for other containers in the system - may contain descriptive information such as, in XML, the validating schema that would be used.
- **GIS (Global Information Section):**
  information that applies to all elements represented in a file, as opposed to information that may pertain to some, but not to all components of the file.
  o Usually contains, for example, the title of the (XML) file, the institution or individual originating the file, address information, copyright information, update information, etc.
- **CI (Complementary Information):**
  o Usually contains, for example, textual bibliographical or administrative information residing in or external to the file, static or dynamic graphic images, video, audio, or virtually any other kind of binary data [i.e. blobs]. Might also include references to other terminological resources or contextual links to related text corpora or to ontologies. These items are often designated as shared resources because they are available to all points in a termbase and are not repeated for different entries.
- **TE (Terminological Entry):**
  information that pertains to a single concept.
  o Usually contains, for example, the terms assigned to a concept, descriptive information pertinent to a concept,

and administrative information concerning the concept. Can contain one or more language sections depending on whether the termbase is monolingual, bilingual, or multilingual.

- **LS (Language Section):**
  contains all the terminological sections (TS) for a terminological entry that are used in a given language, as well as information
  o Usually contains, for example, definitions, contexts, etc. associated with that language or the terms in that language.
- **TS (Term Section):**
  information about terms
  o Usually contains, for example, a single term used to designate the concept that is the subject of the terminological entry, as well as any other information (e.g., definitions, contexts, etc), associated with that term.
- **TCS (Term Component Section):**
  information about morphemic elements, words, or contiguous strings from which a polynomial term is formed. In some languages, such as German or English, it is frequently unnecessary to distinguish information about the individual components making up a polynomial term. In other languages, such as French or Spanish, it is important to be able to include information such as gender for the individual words used in constructing a multiword term because this information is necessary when using the term in texts.

These instances of Structural Levels implement the "dependency" relation with constraints on cardinality (cf. figure 6), which can also be schematized by the sketch shown in figure 7.

- A Terminological Entry can contain any number of Language Sections (0..*);
- A Language Section can contain any number of Term Sections (0..*);
- A Term Section can contain any number of Term Component Sections (0..*)
- A Terminological Data Collection must contain exactly one Global Information block (1..1), at most one Complementary Information block (0..1) and any number of Terminological Entries (0..*);

Hierarchical organization is ensured by the 1..1 limitations expressed for the dual cardinalities for each relation.

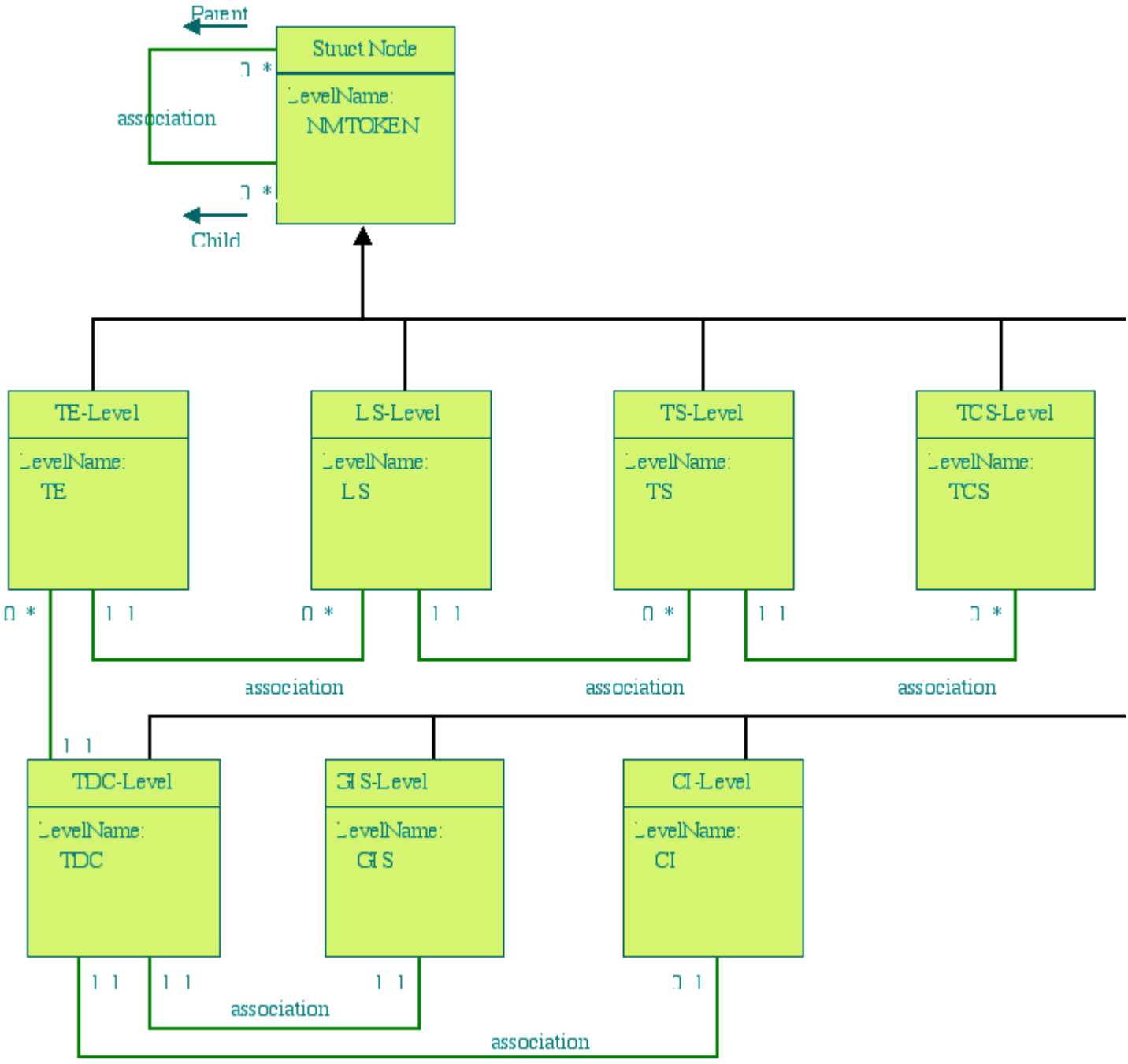



**Figure 6:** **The terminological meta-model.**

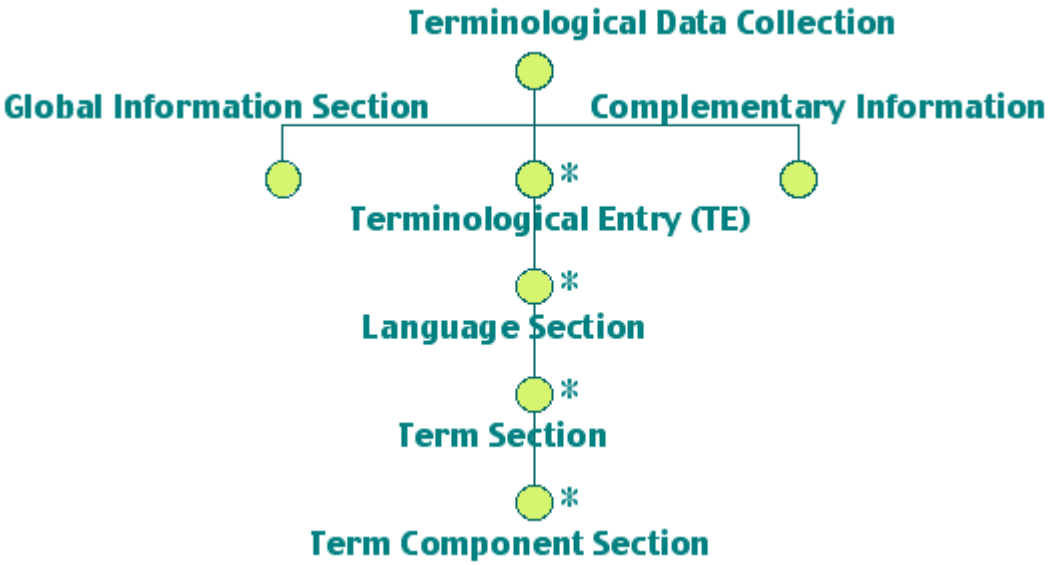



**Figure 7:**

**5.4 Designing representations of terminological data on the basis of the meta-model**

Each data category specification **shall** be configured in response to the real-world needs of some user group and consists of a list of data categories from the metadata registry (ISO 12620) and constraints on each of those data categories. Constraints include restrictions on the values each data category **can** take (ranging from "text with markup" for contextual examples to a "picklist" for grammatical gender, including specific data types, for example, defined in XML Schemas part 2, which **can** be used to describe numbers or dates, for instance). Constraints on descriptive data categories also include restrictions on where a particular data category **can** appear in an entry, selected from the options provided by the core structure of the meta-model (namely, TE, LS, TS and TCS).

The meta-model provides for a vast range of possible data models, particularised to a real-world application by selecting a structure based on the meta-model and restricted to certain data categories and data category values applied to the entry, language section, term section, and term component section as object classes of the meta-model.

All ISO 16642 conformant XML formats are based on (1) the meta-model, (2) subsets of data category specifications that are essentially derived from ISO 12620, and (3) XML DTDs or XML schemas. Database applications for terminological data **can** also differ, but to conform to ISO 16642, they must be based on (1) the meta-model, (2) data category specifications that are essentially based on ISO 12620, and (3) data models defined by means of entity relationship diagrams.

**5.5 Interchange, dissemination and interoperability**

Interchange involves a transfer of information between two computer systems and is often bi-directional, but **need not** be. Interchange is accomplished using intermediate formats. Dissemination is uni-directional and **can** be either for use by another computer system or for human viewing. Formats structured in compliance with the meta-model and the data category specifications of ISO 12620 **should** be interoperable, that is, it **should** be possible to convert data from one format into another format and back without loss of information [sometimes referred to as a "lossless round trip"].

If data are so rigorously defined that it is unnecessary for the importer to establish contact with the originators of the data in order to interpret them, interchange is said to be blind. When there are only two interchange partners and they are known to each other, blindness is not an issue. But when there are multiple sources of terminological data that must be imported by a single routine, especially if it is desirable to add more sources without modifying the import routine, blindness becomes very important.

In bi-directional interchange, the objective is usually to maximise the preservation of information. But in the case of dissemination, representation **can** be intentionally partial, leaving out some information that was in the original data collection. For example, a dissemination-oriented representation for human translators does not necessarily include some administrative information that is only relevant to terminologists maintaining the database.

The specifics of a particular XML format will be influenced by the purpose of the format (for dissemination or for interchange) and by whether there is a need for blindness. Whatever the purposes and real-world needs that guide the design of a database application or an XML format, once designed, the database structure or the format takes on a life of its own as it is used to represent a variety of data, some of which **may** be unanticipated by the designer. By following the integrated approach described here, it is more likely that the resulting format will be adaptable to varied circumstances and that it will be compatible with other database structures or formats that can be shown to be conformant to this

standard.

#### 5.6 XML canonical representation of the generic model

##### 5.6.1 Introduction

The hierarchical organisation of the meta-model and the qualification of each structural level **can**be realised in XML by instantiating the abstract structure shown above and associating information units to this structure.

The meta-model **can** be represented by means of a generic element <struct> (for *structure*) which **can** recursively express the embedding of the various representation levels of a terminological data collection. Each structural node in the meta-model **shall** be identified by means of a *type* attribute associated with the <struct> element. The possible values of the *type* attribute **shall** be the identifiers of the levels in the meta-model, i.e., TDC, GIS, TE, CI, LS, TS, TCS. Basic information units associated with a structural skeleton **can** be represented using the <feat> (for *feature*) element.

Compound information units **can** be represented using the <brack> (for *bracket*) element, which **can** itself contain a <feat> element followed by any combination of <feat> elements and <brack> elements. Each information unit **must** be qualified with a *type* attribute, which **shall** take as its value the name of an ISO 12620 data category or that of a user-defined data category. Finally, the content model of the <feat> element **can** contain annotations expressed by means of an <annot> (for *annotation*) element. This element is also qualified by a *type* attribute referencing an ISO 12620 data category or an equivalent user-defined data category.

##### 5.6.2 Example

The following example illustrates how the information contained within the MSC terminological entry presented in section 5.1.2 **can** be encoded in the GMT format. This entry only contains basic information units and, as seen before, only maps two levels of the meta-model, namely TE (Terminological Entry) and LS (Language Section).

```
<?xml version="1.0" encoding="iso-8859-1"?>

<tmf>
     <struct type="TE">
          <feat type="entry identifier">ID67</feat>
          <feat type="subject field">manufacturing</feat>
          <feat type="definition">A value between 0 and 1 used in ...</feat>
          <struct type="LS">
               <feat type="language identifier">en</feat>
               <struct type="TS">
                    <feat type="term">alpha smoothing factor</feat>
                    <feat type="term type">fullForm</feat>
               </struct>
          </struct>
               <struct type="LS">
                    <feat type="language identifier">hu</feat>
                    <struct type="TS">
                    <feat type="term">Alfa ...</feat>
               </struct>
          </struct>
     </struct>
</tmf>
```

##### 5.6.3 A thorough description of the GMT format

###### 5.6.3.a) The *tmf* element

**Description**

The *tmf* element is the root element for any valid GMT document. It contains both the global information that corresponds to a terminological data collection, the collection itself, and the complementary information comprising external resources in particular, which are needed for describing the various terminological entries.

**Content model (DTD)**

```
<!ELEMENT tmf (struct)>
```

###### 5.6.3.b) The *struct* element

**Description**

The *struct* element **can** be used to represent a structural node in a given structural skeleton. One such structural node will be represented by exactly one *struct* element. The *struct* element is recursive and **may** also contain *feat* and/or *brack* elements to express information units belonging to the corresponding level of the meta-model.

The *struct* element accepts the following XML attributes:

- **type:** categorises the *struct*element by identifying a structural node in the meta-model (TDC, GIS, CI, TE, LS, TS or TCS) represented by the element.
- **target:** a pointer to another *struct* element where is provided the actual content of the current one. The semantics of this pointer is that of a replacement of the current struct by the one which is pointed to.
- **id:** allows to uniquely identify the corresponding *information unit* in the structural skeleton.

**Content model (DTD)**

```
<!ELEMENT struct ((feat|brack)*, struct*)>

<!ATTLIST struct
     type (TDC|GIS|TE|CI|LS|TS|TCS) #REQUIRED
     id ID #IMPLIED
     target CDATA #IMPLIED>
```

#### 5.6.3.c) The *feat* element

**Description**

The *feat* element **can** be used to express any information unit that is either directly attached to a structural node in the structural skeleton (represented by a *struct* element), or grouped together with other information units within a *brack* element. The *feat* element **can** contain any type of tagged data, whether it corresponds to strict annotation by means of the *annot* element or whether a module involving external markup is used for a specific application. In the latter case, the elements from this module **should** be qualified according to a specific namespace.

The *feat* element accepts the following XML attributes:

- **type:** categorises the *feat* element through the reference to the name of the corresponding data category. The name **should** be taken either from ISO 12620 or from a data category described in the context of a specific application and not present in ISO 12620;
- **target :** a *pointer* to a *struct* element in the case the information unit expresses a relation between the current structural node and another structural node in the structural skeleton;

- **source:** a *pointer* to a *struct* element in cases where the information unit is described external to the structural node to which it is supposed to be attached. This approach **can** be used, for instance, to describe a database of conceptual links external to a terminological data collection proper.

**Content model (DTD)**

```
<!ELEMENT feat (#PCDATA | annot)*>

<!ATTLIST feat
     type CDATA #REQUIRED
     target CDATA #IMPLIED
     source CDATA #IMPLIED>
```

**Example:**

The following elements constitute valid expressions of information units:

- Basic information unit attached directly to a structural node (level TE):

```
<struct type="TE">
     <feat type="entry identifier">ID67</feat>
</struct>
```

- Basic information unit whose value is a reference to a structural node in the structural skeleton and whose *id* attribute has the value "TE24":

```
<struct type="TE">
     <feat type="partitive relation" target="TE24"/>
</struct>
```

- Basic information unit anchored at the structural node in the structural skeleton whose *id* attribute value is "TE24":

```
<struct type="TE">
     <feat type="entry identifier" source="TE24">ID67</feat>
</struct>
```

- Compound information unit anchored at the structural node in the structural skeleton whose *id* attribute value is "TE23" and which makes reference to a structural node in the structural skeleton whose "*id*" attribute value is "TE24":

```
<feat type="partitive relation" source="TE23" target="TE24"/>
```

#### 5.6.3.d) The *brack* element

**Description**

The *brack* element **can** be used to express any group of information units whose meaning is interrelated. It contains a list comprising at least one *feat* element followed by any combination of *brack* and *feat* elements. The *brack* element accepts the following attributes:

- **source :** a *pointer* to a *struct* element in cases where the group of information units involved is described external to the structural node to which it is supposed to be attached. The *source* attribute is thus inherited by all the *feat* elements contained in the current *brack* element.

**Content model (DTD)**

```
<!ELEMENT brack (feat+, (feat|brack)*)>

<!ATTLIST brack
     source CDATA #IMPLIED>
```

**Example:**

- Compound information unit comprising two basic features:

```
<brack>
     <feat type="classification code">xxx</feat>
     <feat type="classification system">Lenoch</feat>
</brack>
```

- Compound information unit anchored at a structural node in the structural skeleton where the *id* attribute value is "TE24":

```
<brack source="TE24">
     <feat type="classification code">xxx</feat>
     <feat type="classification system">Lenoch</feat>
</brack>
```

#### 5.6.3.e) The *annot* element

**Description**

The *annot* element **shall** be used to tag any portion of the content of a given *feat* element, provided this procedure is allowed by the content type of the corresponding data category. The *annot* element accepts the following attributes:

- **type :** categorises the *annot* element by referencing the name of the corresponding data category. The name **should** be taken either from ISO 12620 or from a data category specified in the context of a specific application

and not present in ISO 12620;
- **target :** a *pointer* to a *struct* element in cases where the annotation expresses a relation between the current information unit and another structural node in the structural skeleton.

**Content model (DTD)**

```
<!ELEMENT annot (#PCDATA)>

<!ATTLIST annot

type CDATA #REQUIRED

target CDATA #IMPLIED>
```

**Example**

```
<feat type="definition">
     <annot type="broader concept generic">pencil</annot> whose <annot
     type="characteristic">casing</annot> is fixed around a central <annot
     type="characteristic">graphite</annot> medium which is <annot
     type="characteristic">used for writing or making marks</annot>
</feat>
```

#### 5.7 Representing languages in a terminological data collection

Any terminological data collection conformant with TMF **should** clearly distinguish between working language and object language, which are the two types of language information that **can** be attached to any level of the collection.

The working language is the language used to describe any given textual content in the data collection. This information **should** be represented using the xml:lang attribute as defined in the XML recommendation of the W3C and used accordingly. In particular, the scope of the working language is the whole sub-document starting from the element where the information has been declared, unless it is superseded by another working language declaration for some element in this sub-document.

**Note:** As of March 2001, the current recommendation is to follow RFC 3066 (see http://www.ietf.org/rfc/rfc3066.txt), which renders RFC 1766 obsolete.

The object language is the language of the terminological information which is being described at some level in the terminological data collection (typically at the Language Section level). As such, it is represented in TMF as a data category ("language identifier" in ISO 12620) and may be represented in a given TML using any style among those described in this standard. Its possible values are those allowed by the reference data category in ISO 12620 or a reduced set defined for a given TML.

The following example shows how the two types of languages **can** be used within a Language Section expressed in GMT :

```
<struct type="LS" xml:lang="fr">
     <feat type="language identifier">en</feat>
     <feat type="definition">Une valeur entre 0 et 1 utilisée...</feat>
     <struct type="TS">
          <feat type="term" xml:lang="en">alpha smoothing factor</feat>
          <feat type="term type">fullForm</feat>
     </struct>
</struct>
```

This same example **can** be represented in MSC as follows:

```
<langSet lang="en" xml:lang="fr">
     <descrip type="definition">Une valeur entre 0 et 1 utilisée...</descrip>
     <tig>
          <term xml:lang="en">alpha smoothing factor</term>
          <termNote type="termType">fullForm</termNote>
     </tig>
</langSet>
```

and in Geneter :

```
<languageGrp value="en" xml:lang="fr">
     <Definition>Une valeur entre 0 et 1 utilisée...</Definition>
     <termGrp>
          <Term formType='fullForm'>alpha smoothing factor</Term>
     </termGrp>
</languageGrp>
```

## 6 Defining a terminological markup language (TML)

### 6.1 General

The definition of a terminological markup language (or TML) **shall** be considered as a sequence consisting of two different phases:

- A first phase consists of describing those data categories required for this TML, i.e. a Data Category Specification (DCS). This **can** be done by selecting a subset of the data categories specified in ISO 12620 and, if necessary, specifying additional data categories that are needed for the current TML but do not belong to ISO 12620. This phase **shall** lead to the definition of interoperability conditions needed for interaction with other TMLs;
- A second phase corresponds to the realization of the TML as an XML format. This is achieved by providing expansion trees associated with the different structural nodes in the meta-model and by instantiating the necessary information styles and vocabularies required for the data categories that occur in those trees. This phase provides the minimal information to fully specify the XML schemas controlling the valid instances of the TML as well as the filters which **can** transform a TML instance into an GMT instance and vice versa.

These two steps are more precisely defined in the following sections.

### 6.2 Defining interoperability conditions

The definition of interoperability conditions is based upon the specification of the set of data categories that are valid for a given TML. This specification relies upon the provision, for each data category, of a set of properties, which can be modelled as an RDF (Resource Description Framework) representation. These properties are:

- a namespace (in the sense of XML namespaces) that is either the namespace of a Data Category Registry (e.g. ISO 12620) from which the data category is taken or a local namespace associated with the application defined data categories;
- a unique name (DCName property in RDF) within the namespace;
- a type (DCType) which indicates whether the data category describes a possible information unit for the TML (DCType='complex') or is one possible value of an information unit (DCType='simple');
- the list of possible structural nodes (DCLevel) where the data category may occur for the TML;
- the list of the values (Content) that are allowed for the category in the case of a complex data category.

If a data category has been selected from a DCR, the following constraints apply:

- The content description for the data-category is subsumed by the one in ISO 12620. For instance, if the content is defined by a data type, it **should** be a sub-type of the one described in ISO 12620, or, if the content is described as a picklist, it **should** be a subset of the corresponding picklist in ISO 12620;
- The category **can** be applied to a list of structural nodes, which is a subset of the list of authorised structural nodes expressed in ISO 12620.

### 6.3 Implementing a TML

#### 6.3.1 Introduction

Realising interoperability conditions as a TML requires the definition of the XML structures that **can** be used in instances that will describe the corresponding terminological data collection. This requires an *XML outline*, which is the set of XML elements that will implement the structural skeleton of the TML and *anchoring* mechanisms for the various information units that are described in the DCS.

#### 6.3.2 Expanding the structural skeleton (*XML outline* of a TML)

The XML outline of a TML **shall** be defined by associating an XML sub-tree or expansion tree with each structural node in the structural skeleton. For each structural node having a parent in the structural skeleton (that is, for which there exists a higher level in the meta-model) an anchor **shall** also be defined which comprises a node in its parent's expansion tree and to which its own expansion tree **can** be attached.

The XML outline of a TML comprises all the expansion trees associated with its structural skeleton.

#### 6.3.3 Anchoring data categories on the TML XML outline

The expansion tree associated with a structural node consists of a set of XML element nodes. Each of these nodes is a potential *anchor* for the implementation of an information unit that is allowed at that structural node. According to the anchoring style associated with the information unit and in conjunction with the corresponding vocabulary used in the actual TML, it is possible to specify how the data category will be expressed as an XML sub-structure of its anchor. The corresponding properties (anchor, style, vocabulary) **shall** be additional descriptions included with the corresponding data category in the full DCS associated with a TML.

##### 6.3.3.a) Styles and vocabulary

Any information unit attached to the structural skeleton of a TML **can** be implemented using one of the five styles Attribute, Element, TypedElement, ValuedElement and TypedValuedElement. These styles correspond to the way a feature-value pair is expressed in XML.

The Attribute style implements an information unit as an XML attribute of a given anchor. The vocabulary represents the name of the XML attribute. The value associated with a specific information unit is realised as the content of this XML attribute.

**Examples:**

| GMT representation | MSC representation (anchor: <termEntry>) |
|---|---|
| <struct type="TE"><br><br><feat type="entry identifier">ID67</feat><br><br>...<br><br></struct> | <termEntry id="ID67"><br><br>...<br><br></termEntry> |

| GMT representation | Geneter representation (anchor: <ldl>) |
|---|---|
| <struct type="LS"><br><br><feat type="language identifier">en</feat><br><br>...<br><br></struct> | <ldl language ="en"><br><br>...<br><br></ldl> |

The Element style implements an information unit as an XML element, which itself is a child of a given anchor. The vocabulary represents the name of this XML element. The value associated with a specific information unit is realised as the content of this XML element.

**Examples:**

| GMT representation | MSC representation (anchor: <tig>) |
|---|---|
| <struct type="TS"><br><br><feat type="term">alpha smoothing factor</feat><br><br></struct> | <tig><br><br><term>alpha smoothing factor</term><br><br></tig> |

| GMT representation | Geneter representation (anchor: <tl>) |
|---|---|
| <struct type="TS"> | <tl> |

| | |
|---|---|
| <feat type="term">barbed wire</feat><br><br></struct> | <Term>barbed wire</Term><br><br></tl> |

The TypedElement style implements an information unit as an XML element, which itself is a child of a given anchor, and which is further specified by an XML attribute *type*. The vocabulary represents the name of this XML element and a value for the XML attribute *type*. The value associated with a specific information unit is realised as the content of this XML element.

**Examples:**

| GMT representation | MSC representation (anchor: <termEntry>) |
|---|---|
| <struct type="TE"><br><br><feat type="subject field">manufacturing</feat><br><br></struct> | <termEntry ... ><br><br><descrip type="subject field">manufacturing</descrip><br><br></termEntry> |

| GMT representation | Geneter representation |
|---|---|
| <struct type="TE"><br><br><feat type="subject field">manufacturing</feat><br><br></struct> | <terminologicalEntry ... ><br><br><free type="subject field">manufacturing</free><br><br></terminologicalEntry> |

The ValuedElement style implements an information unit as an XML element, which is itself a child of a given anchor, and which is further specifed by an XML attribute - *value*. The vocabulary represents the name of this XML element. The value associated with a specific information unit is realised as the content of the XML attribute *value*.

The TypedValuedElement style implements an information unit as an XML element, which is itself a child of a given anchor, and which is further specified by means of an XML attribute *type*. The vocabulary represents the name of this XML element and the name of the XML attribute. Thevalue associated with a specific information unit is realised as the content of the XML attribute *type*.

#### 6.3.3.b) Constraints on datatypes for information units

Whereas information units implemented using the Element or TypedElement styles **can** take values that **can** contain additional markup (in particular markup resulting from the implementation of annotations), information units implemented using either the Attribute or ValuedElement styles **shall not** contain any such markup. In this respect, consistency checking is required when defining the DCS for a given TML.

#### 6.3.3.c) External markup modules

DCSs **can** include reference to external markup modules that **can** be used in the content model of information units or to reference external objects corresponding to complementary information in the meta-model (e.g., bibliographical references). These modules are referenced using a registered namespace.

As an example, the following XML schema declaration **can** be used to define a content model comprising any element coming from the xhtml recommendation.

```
<complexType name="xhtmlContent">
     <any namespace="http://www.w3.org/1999/xhtml"
     minOccurs="0"
     maxOccurs="unbounded"
     processContents="skip"/>
</complexType>
```

#### 6.3.4 Implementing annotations

Annotations (expressed in GMT with an <annot> element) **must** be implemented in the same way as information units that are attached to the structural skeleton of a given TML except that only the Element and TypedElement styles **may** be used for this purpose.

#### 6.3.5 Implementing brackets

Bracketed information units (expressed in GMT with a <brack> element) **must** be implemented by providing the name of an element, which is associated to the main information unit of the group and whose content is the set of the realizations of the information units of the group.

#### 6.3.6 Namespaces

The DCS **may** comprise the description of an XML namespace that references the various XML objects (element or attributes) resulting from the definition of an actual TML. This description becomes mandatory when additional markup modules are associated with the TML.

## Annex A (normative): XML schema of the GMT format

The present annex comprises the definition of the GMT format using the XML schema syntax, as of the May 2nd, 2001 recommendation. This schema **should** be used as a reference to check the conformity of any data represented in GMT in the case it does not contain any additional markup module. In any other case, the schema **should** be modified to incorporate the definition of the namespaces to be associated with the external markup to be used.

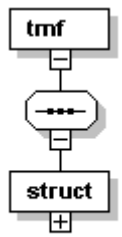
### Schema **gmt.xsd**

Elements Complex types

**tmf**

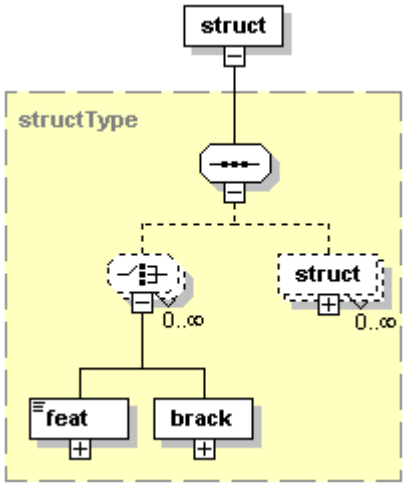
**annotType**
**brackType**
**featType**
**structType**

element **tmf**

| | |
|---|---|
| diagram |   |
| children | **struct** |
| source | <xs:element name="tmf"><br><xs:complexType><br><xs:sequence><br><xs:element name="struct" type="structType"/><br></xs:sequence><br></xs:complexType><br></xs:element> |

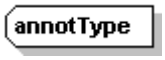
element **tmf/struct**

| | |
|---|---|
| diagram | |

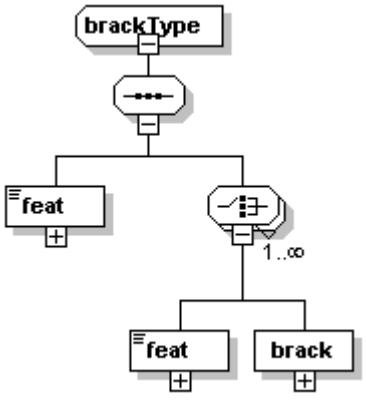



| | |
|---|---|
| type | **structType** |
| children | **featbrackstruct** |
| attributes | Name Type Use Default Fixed<br>type xs:NMTOKEN required |
| source | <xs:element name="struct" type="structType"/> |

complexType **annotType**

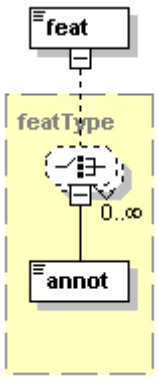
| | |
|---|---|
| diagram |  |
| type | restriction of **xs:string** |
| used by | element **featType/annot** |
| source | <xs:complexType name="annotType"><br><xs:simpleContent><br><xs:restriction base="xs:string"><br><xs:attribute name="type" type="xs:string"/><br><xs:attribute name="target" type="xs:string"/><br></xs:restriction><br></xs:simpleContent><br></xs:complexType> |

complexType **brackType**

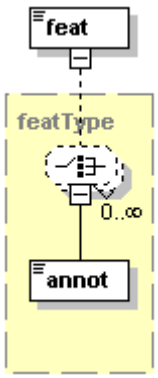
| | |
|---|---|
| diagram |  |
| children | **featfeatbrack** |
| used by | elements **brackType/brackstructType/brack** |
| attributes | Name Type Use Default Fixed<br>source xs:string |
| source | <xs:complexType name="brackType"><br><xs:sequence><br><xs:element name="feat" type="featType"/><br><xs:choice maxOccurs="unbounded"><br><xs:element name="feat" type="featType"/><br><xs:element name="brack" type="brackType"/> |

```
    </xs:choice>
  </xs:sequence>
  <xs:attribute name="source" type="xs:string"/>
</xs:complexType>
```

element **brackType/feat**

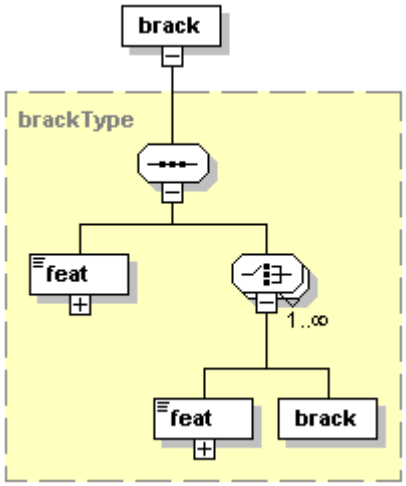
| | |
|---|---|
| diagram |   |
| type | **featType** |
| children | **annot** |
| attributes | Name Type Use Default Fixed<br>type xs:string required<br>target xs:string required<br>source xs:string |
| source | <xs:element name="feat" type="featType"/> |

element **brackType/feat**

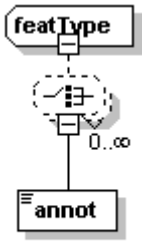
| | |
|---|---|
| diagram |   |
| type | **featType** |
| children | **annot** |
| attributes | Name Type Use Default Fixed<br>type xs:string required<br>target xs:string required<br>source xs:string |
| source | <xs:element name="feat" type="featType"/> |

element **brackType/brack**

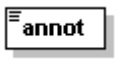
| | |
|---|---|
| diagram | |

| | |
|---|---|
| |  |
| type | **brackType** |
| children | **featfeatbrack** |
| attributes | Name Type Use Default Fixed<br>source xs:string |
| source | <xs:element name="brack" type="brackType"/> |

complexType **featType**

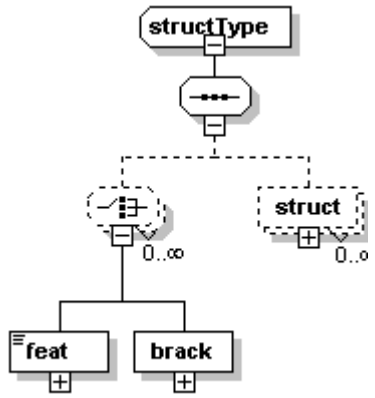
| | |
|---|---|
| diagram |  |
| children | **annot** |
| used by | elements **brackType/featbrackType/featstructType/feat** |
| attributes | Name Type Use Default Fixed<br>type xs:string required<br>target xs:string required<br>source xs:string |
| source | <xs:complexType name="featType" mixed="true"><br><xs:choice minOccurs="0" maxOccurs="unbounded"><br><xs:element name="annot" type="annotType"/><br></xs:choice><br><xs:attribute name="type" type="xs:string" use="required"/><br><xs:attribute name="target" type="xs:string" use="required"/><br><xs:attribute name="source" type="xs:string"/><br></xs:complexType> |

element **featType/annot**

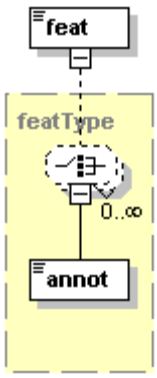
| | |
|---|---|
| diagram |  |
| type | **annotType** |
| attributes | Name Type Use Default Fixed<br>type xs:string<br>target xs:string |
| source | <xs:element name="annot" type="annotType"/> |

complexType **structType**

| | |
|---|---|
| diagram | |

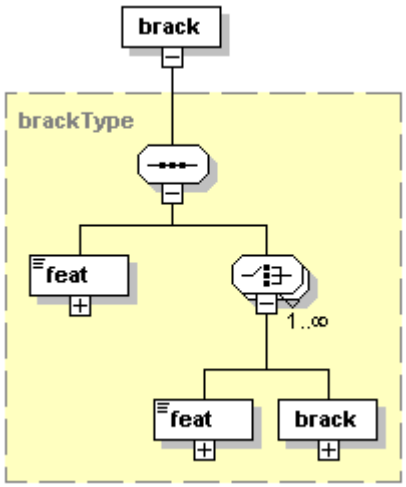
| | |
|---|---|
| |  |
| children | **featbrackstruct** |
| used by | elements **tmf/structstructType/struct** |
| attributes | Name Type Use Default Fixed<br>type xs:NMTOKEN required |
| source | <xs:complexType name="structType"><br><xs:sequence><br><xs:choice minOccurs="0" maxOccurs="unbounded"><br><xs:element name="feat" type="featType"/><br><xs:element name="brack" type="brackType"/><br></xs:choice><br><xs:element name="struct" type="structType" minOccurs="0" maxOccurs="unbounded"/><br></xs:sequence><br><xs:attribute name="type" use="required"><br><xs:simpleType><br><xs:restriction base="xs:NMTOKEN"><br><xs:enumeration value="TDC"/><br><xs:enumeration value="GIS"/><br><xs:enumeration value="CI"/><br><xs:enumeration value="TE"/><br><xs:enumeration value="LS"/><br><xs:enumeration value="TS"/><br><xs:enumeration value="TCS"/><br></xs:restriction><br></xs:simpleType><br></xs:attribute><br></xs:complexType> |

element **structType/feat**

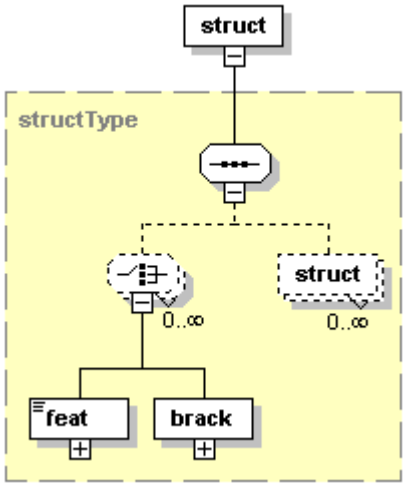
| | |
|---|---|
| diagram |  |
| type | **featType** |
| children | **annot** |
| attributes | Name Type Use Default Fixed<br>type xs:string required<br>target xs:string required<br>source xs:string |
| source | <xs:element name="feat" type="featType"/> |

element **structType/brack**

| | |
|---|---|
| diagram | |

| | |
|---|---|
| |  |
| type | **brackType** |
| children | **featfeatbrack** |
| attributes | Name Type Use Default Fixed<br>source xs:string |
| source | <xs:element name="brack" type="brackType"/> |

element **structType/struct**

| | |
|---|---|
| diagram |  |
| type | **structType** |
| children | **featbrackstruct** |
| attributes | Name Type Use Default Fixed<br>type xs:NMTOKEN required |
| source | <xs:element name="struct" type="structType" minOccurs="0" maxOccurs="unbounded"/> |

```
<?xml version="1.0" encoding="UTF-8"?>

<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
     <xs:complexType name="annotType">
          <xs:simpleContent>
               <xs:restriction base="xs:string">
                    <xs:attribute name="type" type="xs:string"/>
                    <xs:attribute name="target" type="xs:string"/>
               </xs:restriction>
          </xs:simpleContent>
     </xs:complexType>
     <xs:complexType name="brackType">
          <xs:sequence>
               <xs:element name="feat" type="featType"/>
               <xs:choice maxOccurs="unbounded">
```

```
                    <xs:element name="feat" type="featType"/>
                    <xs:element name="brack" type="brackType"/>
               </xs:choice>
          </xs:sequence>
          <xs:attribute name="source" type="xs:string"/>
     </xs:complexType>
     <xs:complexType name="featType" mixed="true">
          <xs:choice minOccurs="0" maxOccurs="unbounded">
               <xs:element name="annot" type="annotType"/>
          </xs:choice>
          <xs:attribute name="type" type="xs:string" use="required"/>
          <xs:attribute name="target" type="xs:string" use="required"/>
          <xs:attribute name="source" type="xs:string"/>
     </xs:complexType>
     <xs:complexType name="structType">
          <xs:sequence>
               <xs:choice minOccurs="0" maxOccurs="unbounded">
                    <xs:element name="feat" type="featType"/>
                    <xs:element name="brack" type="brackType"/>
               </xs:choice>
               <xs:element name="struct" type="structType" minOccurs="0"
               maxOccurs="unbounded"/>
          </xs:sequence>
          <xs:attribute name="type" use="required">
               <xs:simpleType>
                    <xs:restriction base="xs:NMTOKEN">
                         <xs:enumeration value="TDC"/>
                         <xs:enumeration value="GIS"/>
                         <xs:enumeration value="CI"/>
                         <xs:enumeration value="TE"/>
                         <xs:enumeration value="LS"/>
                         <xs:enumeration value="TS"/>
                         <xs:enumeration value="TCS"/>
                    </xs:restriction>
               </xs:simpleType>
          </xs:attribute>
     </xs:complexType>
     <xs:element name="tmf">
          <xs:complexType>
               <xs:sequence>
                    <xs:element name="struct" type="structType"/>
               </xs:sequence>
          </xs:complexType>
     </xs:element>
</xs:schema>
```

## Annex B (normative): The MSC Terminology Markup Language

### B.1 Introduction

A TML is defined by specifying a set of data categories and an XML implementation of the meta-model elaborated in this International Standard combined with these data categories. (This meta-model is hereinafter simply referred to as "the meta-model".) The TML called MSC (MARTIF with Specified Constraints) is defined by a specified set of data categories selected from ISO 12620.

MSC is designed to represent terminological data for the processes of analysis, dissemination, and exchange of information from human-oriented terminological databases (termbases).

The data categories in MSC have been selected with the objective of supporting "blind" representation. A blind representation supports interpretation without consulting the data provider. For example, consider the data category *grammatical gender*. Suppose that a TML specifies that the possible values of this data category are *masculine*, *feminine*, and *neuter*, with their normal meanings in grammars of Western European languages. An instance of the data category grammatical *gender* with the value *feminine* can be interpreted without consulting the data provider. All information necessary to the interpretation of this data category is in the definition of the TML. On the other hand, suppose that another TML does not specify the possible values of the data category *grammatical gender*. Instead, these values are to be determined separately by each data provider. A user who receives a terminological data collection including the data category *grammatical gender* with the value *356* will not be able to interpret this data category without consulting the data provider concerning the meaning of the value *356*. Blindness is not an absolute property of a format but is instead a

matter of degree.

The implementation of the meta-model in MSC is based on ISO 12200, which is commonly called MARTIF. Each implementation of the meta-model specifies expansion trees for every structural node of the meta-model and, for each data category, an XML style and vocabulary. For example, the structural node *language section* is expanded in MSC as the XML element *langSet* . And the data category *definition* uses the typedElement style with the vocabulary *descrip* for the tag name and *definition* for the value of the type attribute (`<descrip type="definition">...</descrip>`). This implementation has been chosen in order to support multiple modes of compliance checking. All TMLs support the mode of compliance checking that uses a comprehensive XML schema for the TML as input to a general-purpose XML parser. This mode produces error messages suitable for an expert user of MSC who is thoroughly familiar with both XML and the comprehensive MSC schema. The extensive use in MSC of typedElement styles also makes it possible to report MSC compliance of an XML document using a two-stage rather than a single-stage mode. In the first stage, an MSC XML document is checked against the proper use of typed elements and structural nodes; in the second stage, the document is checked against specific data categories. This alternative mode of compliance checking supports the generation of error messages that are more readily understandable to a terminologist with a limited familiarity with XML and no experience in reading XML schemas.

Any number of additional TMLs, besides MSC, can be defined using the same meta-model implementation as MSC together with a different selection of data categories. Any given set can be either more limited than the data categories of MSC (i.e., constitute a subset of MSC) or more extensive (i.e., constitute a superset), depending on the application. A terminologist-friendly compliance checker can be designed to adjust automatically to the particular set of data categories chosen for that TML.

#### B.2 An example of an MSC XML document

The following is an example of a simple but complete MSC document. The numbers in square brackets to the left of certain lines are not part of the MSC document. They serve as footnote numbers to the comments below.

**Note:** a sample MSC document is available in machine-readable form; the following is only illustrative.

```
[1] <?xml version='1.0'?>
    <!DOCTYPE martif SYSTEM "./MSCcoreStructureDTD-v-1-0.DTD.TXT">
[2] <martif type='MSC' xml:lang='en' >
[3]       <martifHeader>
               <fileDesc><sourceDesc><p>from an Oracle corporation
               termBase</p></sourceDesc></fileDesc>
               <encodingDesc><p type='DCSName'>MSCdefaultXCS-v-1-
               0.XML</p></encodingDesc>
          </martifHeader>
[4]       <text> <body>
[5]            <termEntry id='eid-Oracle-67'>
[6]                 <descrip type='subjectField'>manufacturing</descrip>
[7]                 <descrip type='definition'>A value between 0 and 1 used
                    in ...</descrip>
[8]                 <langSet xml:ang='en'>
[9]                      <tig>
                              <term tid='tid-Oracle-67-en1'>alpha smoothing
                              factor</term>
[10]                          <termNote type='termType'>fullForm</termNote>
[11]                     </tig>
[12]                </langSet>
[13]                <langSet xml:lang='hu'>
[14]                     <tig>
                              <term tid='tid-Oracle-67-hu1'>Alfa simítási tényezõ
                              </term>
                         </tig>
[15]                </langSet>
[16]           </termEntry>
[17]      </body> </text>
[18] </martif>
```

Only a minimal acquaintance with XML is assumed in the following explanation. Indeed, an acquaintance with HTML from building simple web pages, along with the knowledge that XML allows user-defined tag names whereas HTML comes with a set of pre-defined tag names, should be sufficient to allow an understanding of the following explanation. For key MSC elements, the correspondence to the structural component of the meta-model in ISO 16642 [forthcoming] (the TMF project) is indicated.

[1] <?XML ... : These lines state 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) and to the MSC DTD.

[2] <martif ... : This line states that this particular XML document is a MSC document and thus can be validated against a specification of the MSC core structure, which, for this document, is an XML DTD called MSCdefaultXCS-v-0-1.XML and against the XCS file referred to in the encodingDescription element. Alternatively, the core structure can be validated against a schema version of the description of the core structure. 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, e.g. "fr-CA" for Canadian French.

[3] <martifHeader ... : These lines provide global information about the collection: specifically, a file description indicating that the example was derived from an entry in a termbase used at Oracle corporation and that the MSC XCS (MSC-XCSV04a), not to be confused with the MSC core DTD (MSC-corestrucV04a) is being used.

[4] <text> <body>: The text element surrounds the body element, which contains the collection of concept-oriented "Teminological Entry" (<termEntry>) elements, and, optionally, <front> and <back> elements.

[5] <termEntry ... : Each termEntry element is one instance of the "Terminological Entry" object class as illustrated by the data model 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].

[6] <descrip type='subjectField' ...: The subject field data category is authorized by the XCS (Data Constraint Specification) mentioned above. It consists of a meta data category element (**descrip**) with the specific data category indicated in the value of the type attribute.

[7] <descrip type='definition' ... : This piece of descriptive information is also associated with the concept.

[8] <langSet lang='en'>: The langSet element corresponds to a "Language Section" object class, according to which a Terminological Entry consists of associated information and language sections. This line begins the English Language Section.

[9] <tig><term> ... : The meta-model in ISO 16642 states that a Language Section consists of instances of a "Term Section" object class, which, in MSC 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 *termType*. The name *tig* stands for **t**erm **i**nformation **g**roup. The id 'tid-Oracle-67-en1' consists of the information: eid [entry 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].

[10] <termNote type='termType' ... : This piece of information associated with the term is the 12620 data category "term type". Its value in this case is "fullForm". A *termNote* tag is used since the information is closely associated with the term itself rather than with the concept being described.

[11] </tig>: This element simply ends the current Term Section.

[12] </langSet>: This element ends the English Language Section.

[13] <langSet lang='hu'>: This element begins the Hungarian Language Section.

[14] <tig> ...: This line consists of a Term Section with a Hungarian term but no definition and no explicit term type. Each character of the term that is not found in ISO 646 is represented as a hex character reference corresponding directly to a Unicode character. The actual Hungarian term is "Alfa simítási tényezõ". Note that the final character "õ" (o-tilde) should more properly be an o-double-acute, which is represented by the following Unicode hex character reference: "ő", a character not available in a typical Latin 1 font. In XML, a Unicode hex character reference consists of "&#x" + four hex digits from the Unicode standard + a semicolon.

[15] </langSet>;: This element ends the Hungarian Language Section.

[16] `</termEntry>`: This element ends the current Terminological Entry.

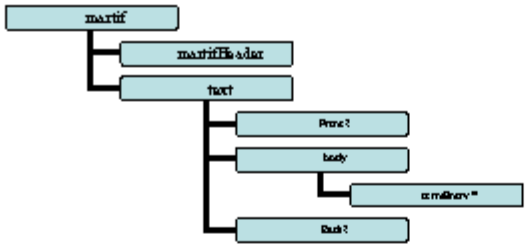
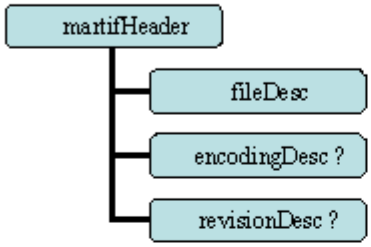
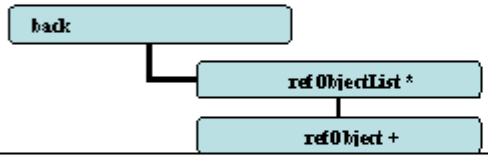
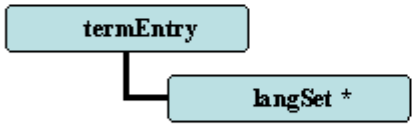
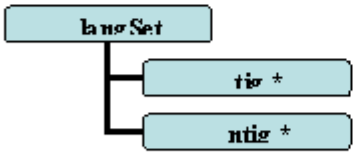
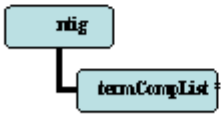
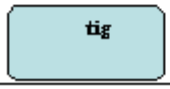
[17] `</body> </text>`: These elements end the set of terminological entries, which in this case consist of only one entry, and the MSC text element, which is the composite of terminological entries and other resources called *complementary information* in the meta-model [see 6.3]. In this MSC document, there are no resources outside the terminological entry. If there were, they would be in the MSC element *back* or *front*.

[18] `</martif>`: This element ends the entire MSC document.

This sample MSC entry has several properties:

- 1. It corresponds directly to the meta-model in ISO 16642 (the TMF project).
- 2. It is well-formed and core-structure-valid.
- 3. It adheres to the default MSC extensible constraint specification (XCS)

**B.3 Expansion trees**

| Level Structure | comments |
|---|---|
| TDC  | The following codes apply to all the structure charts:<br>? occurs 0 or 1 times<br>* repeatable element<br>+ occurs at least once |
| GIS  | There are more elements which are descendants of fileDesc, encodingDesc and revisionDesc which could not be shown in the limited space. |
| CI  | The type values for refObject are expressed in the XCS file but are not datacategories included in ISO 12620. |
| TE  <br>LS  <br>TS  <br>TS  <br>TCS  | ntigs differ from tigs in that they allow for a termComponent level. In this way they are more robust than tigs. |

### B.4 Data categories

| DatCat: "**abbreviated form for**" | |
|---|---|
| **DC identifier:** ISO12620A-02013002<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote abbreviatedFormFor<br>**Example:** `<termNote type="abbreviatedFormFor">DatCat Value</termNote>` |

| DatCat: "**administrative status**" | |
|---|---|
| **DC identifier:** ISO12620A-020903<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote administrativeStatus |

| | |
|---|---|
| | **Example:** `<termNote type="administrativeStatus">DatCat Value</termNote>` |

| DatCat: "**standardized term-admn-sts**" | |
|---|---|
| **DC identifier:** ISO12620A-02090301<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** standardizedTerm-admn-sts<br>**Example:** `standardizedTerm-admn-sts` |

| DatCat: "**preferred term-admn-sts**" | |
|---|---|
| **DC identifier:** ISO12620A-02090302<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** preferredTerm-admn-sts<br>**Example:** `preferredTerm-admn-sts` |

| DatCat: "**admitted term-admn-sts**" | |
|---|---|
| **DC identifier:** ISO12620A-02090303<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** admittedTerm-admn-sts<br>**Example:** `admittedTerm-admn-sts` |

| DatCat: "**deprecated term-admn-sts**" | |
|---|---|
| **DC identifier:** ISO12620A-02090304<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** deprecatedTerm-admn-sts<br>**Example:** `deprecatedTerm-admn-sts` |

| DatCat: "**superseded term-admn-sts**" | |
|---|---|
| **DC identifier:** ISO12620A-02090305<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** supersededTerm-admn-sts<br>**Example:** `supersededTerm-admn-sts` |

| DatCat: "**legal term-admn-sts**" | |
|---|---|
| **DC identifier:** ISO12620A-02090306<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** legalTerm-admn-sts<br>**Example:** `legalTerm-admn-sts` |

| DatCat: "**regulated term-admn-sts**" | |
|---|---|
| **DC identifier:** ISO12620A-02090307<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** regulatedTerm-admn-sts<br>**Example:** `regulatedTerm-admn-sts` |

| DatCat: "**animacy**" | |
|---|---|
| **DC identifier:** ISO12620A-020204<br>**Level(s):** TCS TS | **Style:** TypedElement<br>**Vocabulary:** termNote animacy<br>**Example:** `<termNote type="animacy">DatCat Value</termNote>` |

| DatCat: "**animate**" | |
|---|---|
| **DC identifier:** ISO12620A-02020401<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** animate<br>**Example:** `animate` |

DatCat: **"inanimate"**

**DC identifier:** ISO12620A-02020402
**Level(s):**

**Style:** Simple
**Vocabulary:** inanimate
**Example:** `inanimate`

DatCat: **"other animacy"**

**DC identifier:** ISO12620A-02020403
**Level(s):**

**Style:** Simple
**Vocabulary:** otherAnimacy
**Example:** `otherAnimacy`

DatCat: **"antonym-concept"**

**DC identifier:** ISO12620A-10180602
**Level(s):** TE

**Style:** TypedElement
**Vocabulary:** descrip antonym-concept
**Example:** `<descrip type="antonym-concept">DatCat Value</descrip>`

DatCat: **"antonym-term"**

**DC identifier:** ISO12620A-10180601
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** descrip antonym-term
**Example:** `<descrip type="antonym-term">DatCat Value</descrip>`

DatCat: **"application subset"**

**DC identifier:** ISO12620A-100306
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** admin applicationSubset
**Example:** `<admin type="applicationSubset">DatCat Value</admin>`

DatCat: **"associated concept"**

**DC identifier:** ISO12620A-070210
**Level(s):** LS TE

**Style:** TypedElement
**Vocabulary:** descrip associatedConcept
**Example:** `<descrip type="associatedConcept">DatCat Value</descrip>`

DatCat: **"audio"**

**DC identifier:** ISO12620A-050502
**Level(s):** LS TE TS

**Style:** TypedElement
**Vocabulary:** descrip audio
**Example:** `<descrip type="audio">DatCat Value</descrip>`

DatCat: **"broader concept generic"**

**DC identifier:** ISO12620A-070201
**Level(s):** LS TE

**Style:** TypedElement
**Vocabulary:** descrip broaderConceptGeneric
**Example:** `<descrip type="broaderConceptGeneric">DatCat Value</descrip>`

DatCat: **"broader concept partitive"**

**DC identifier:** ISO12620A-070202
**Level(s):** LS TE

**Style:** TypedElement
**Vocabulary:** descrip broaderConceptPartitive
**Example:** `<descrip type="broaderConceptPartitive">DatCat Value</descrip>`

DatCat: **"business unit subset"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100308<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** admin businessUnitSubset<br>**Example:** `<admin type="businessUnitSubset">DatCat Value</admin>` |

DatCat: **"characteristic"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0508<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** descrip characteristic<br>**Example:** `<descrip type="characteristic">DatCat Value</descrip>` |

DatCat: **"classification code"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0402<br>**Level(s):** LS TE TS | **Style:** TypedElement<br>**Vocabulary:** descrip classificationCode<br>**Example:** `<descrip type="classificationCode">DatCat Value</descrip>` |

DatCat: **"concept identifier"**

| | |
|---|---|
| **DC identifier:** ISO12620A-1014<br>**Level(s):** TE | **Style:** Attribute<br>**Vocabulary:** id<br>**Example:** `id="DatCat Value"` |

DatCat: **"concept origin"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0509<br>**Level(s):** LS TE TS | **Style:** TypedElement<br>**Vocabulary:** admin conceptOrigin<br>**Example:** `<admin type="conceptOrigin">DatCat Value</admin>` |

DatCat: **"concept position"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0702<br>**Level(s):** LS TE | **Style:** TypedElement<br>**Vocabulary:** descrip conceptPosition<br>**Example:** `<descrip type="conceptPosition">DatCat Value</descrip>` |

DatCat: **"context"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0503<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** descrip context<br>**Example:** `<descrip type="context">DatCat Value</descrip>` |

DatCat: **"context type"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0803<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** descripNote contextType<br>**Example:** `<descripNote type="contextType">DatCat Value</descripNote>` |

DatCat: **"defining context"**

| | |
|---|---|
| **DC identifier:** ISO12620A-080301<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** definingContext<br>**Example:** `definingContext` |

DatCat: **"explanatory context"**

| | |
|---|---|
| **DC identifier:** ISO12620A-080302<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** explanatoryContext |

| | |
|---|---|
| | **Example:** `explanatoryContext` |

DatCat: "**associative context**"

| | |
|---|---|
| **DC identifier:** ISO12620A-080303<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** associativeContext<br>**Example:** `associativeContext` |

DatCat: "**linguistic context**"

| | |
|---|---|
| **DC identifier:** ISO12620A-080304<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** linguisticContext<br>**Example:** `linguisticContext` |

DatCat: "**metalinguistic context**"

| | |
|---|---|
| **DC identifier:** ISO12620A-080305<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** metalinguisticContext<br>**Example:** `metalinguisticContext` |

DatCat: "**translated context**"

| | |
|---|---|
| **DC identifier:** ISO12620A-080306<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** translatedContext<br>**Example:** `translatedContext` |

DatCat: "**coordinate concept generic**"

| | |
|---|---|
| **DC identifier:** ISO12620A-07020401<br>**Level(s):** LS TE | **Style:** TypedElement<br>**Vocabulary:** descrip coordinateConceptGeneric<br>**Example:** `<descrip type="coordinateConceptGeneric">`*`DatCat Value`*`</descrip>` |

DatCat: "**coordinate concept partitive**"

| | |
|---|---|
| **DC identifier:** ISO12620A-07020402<br>**Level(s):** LS TE | **Style:** TypedElement<br>**Vocabulary:** descrip coordinateConceptPartitive<br>**Example:** `<descrip type="coordinateConceptPartitive">`*`DatCat Value`*`</descrip>` |

DatCat: "**corpus trace**"

| | |
|---|---|
| **DC identifier:**<br>**Level(s):** | |

DatCat: "**cross-reference**"

| | |
|---|---|
| **DC identifier:** ISO12620A-1018<br>**Level(s):** TE LS TS | **Style:** TypedElement<br>**Vocabulary:** ref cross-reference<br>**Example:** `<ref type="cross-reference">`*`DatCat Value`*`</ref>` |

DatCat: "**customer subset**"

| | |
|---|---|
| **DC identifier:** ISO12620A-100301<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** admin customerSubset<br>**Example:** `<admin type="customerSubset">`*`DatCat Value`*`</admin>` |

DatCat: "**database type**"

**DC identifier:**
**Level(s):**

DatCat: "**date**"

**DC identifier:** ISO12620A-100201
**Level(s):** TE LS TS TCS

**Style:** Element
**Vocabulary:** date
**Example:** `<date>DatCat Value</date>`

DatCat: "**definition**"

**DC identifier:** ISO12620A-0501
**Level(s):** LS TE TS

**Style:** TypedElement
**Vocabulary:** descrip definition
**Example:** `<descrip type="definition">DatCat Value</descrip>`

DatCat: "**definition type**"

**DC identifier:** ISO12620A-0802
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** descripNote definitionType
**Example:** `<descripNote type="definitionType">DatCat Value</descripNote>`

DatCat: "**intensional definition**"

**DC identifier:** ISO12620A-080201
**Level(s):**

**Style:** Simple
**Vocabulary:** intensionalDefinition
**Example:** `intensionalDefinition`

DatCat: "**extensional definition**"

**DC identifier:** ISO12620A-080202
**Level(s):**

**Style:** Simple
**Vocabulary:** extensionalDefinition
**Example:** `extensionalDefinition`

DatCat: "**partitive definition**"

**DC identifier:** ISO12620A-080203
**Level(s):**

**Style:** Simple
**Vocabulary:** partitiveDefinition
**Example:** `partitiveDefinition`

DatCat: "**translated definition**"

**DC identifier:** ISO12620A-080204
**Level(s):**

**Style:** Simple
**Vocabulary:** translatedDefinition
**Example:** `translatedDefinition`

DatCat: "**description type**"

**DC identifier:** ISO12620A-0801
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** descripNote descriptionType
**Example:** `<descripNote type="descriptionType">DatCat Value</descripNote>`

DatCat: "**domain expert**"

**DC identifier:** ISO12620A-1025

**Style:** TypedElement

| | |
|---|---|
| **Level(s):** LS TE TS | **Vocabulary:** admin domainExpert<br>**Example:** `<admin type="domainExpert">DatCat Value</admin>` |

DatCat: **"element working status"**

| | |
|---|---|
| **DC identifier:** ISO12620A-1011<br>**Level(s):** TE LS TS TCS | **Style:** TypedElement<br>**Vocabulary:** admin elementWorkingStatus<br>**Example:** `<admin type="elementWorkingStatus">DatCat Value</admin>` |

DatCat: **"starter element"**

| | |
|---|---|
| **DC identifier:** ISO12620A-101101<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** starterElement<br>**Example:** `starterElement` |

DatCat: **"working element"**

| | |
|---|---|
| **DC identifier:** ISO12620A-101102<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** workingElement<br>**Example:** `workingElement` |

DatCat: **"consolidated element"**

| | |
|---|---|
| **DC identifier:** ISO12620A-101103<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** consolidatedElement<br>**Example:** `consolidatedElement` |

DatCat: **"archive element"**

| | |
|---|---|
| **DC identifier:** ISO12620A-101104<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** archiveElement<br>**Example:** `archiveElement` |

DatCat: **"imported element"**

| | |
|---|---|
| **DC identifier:** ISO12620A-101105<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** importedElement<br>**Example:** `importedElement` |

DatCat: **"exported element"**

| | |
|---|---|
| **DC identifier:** ISO12620A-101106<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** exportedElement<br>**Example:** `exportedElement` |

DatCat: **"entailed term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100601<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** hi entailedTerm<br>**Example:** `<hi type="entailedTerm">DatCat Value</hi>` |

DatCat: **"entry identifier"**

| | |
|---|---|
| **DC identifier:** ISO12620A-1015<br>**Level(s):** TE | **Style:** Attribute<br>**Vocabulary:** id<br>**Example:** `id="DatCat Value"` |

DatCat: **"entry source"**

**DC identifier:** ISO12620A-1013
**Level(s):** TE

**Style:** TypedElement
**Vocabulary:** admin entrySource
**Example:** `<admin type="entrySource">DatCat Value</admin>`

DatCat: **"environment subset"**

**DC identifier:** ISO12620A-100307
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** admin environmentSubset
**Example:** `<admin type="environmentSubset">DatCat Value</admin>`

DatCat: **"etymology"**

**DC identifier:** ISO12620A-020402
**Level(s):** TS TCS

**Style:** TypedElement
**Vocabulary:** termNote etymology
**Example:** `<termNote type="etymology">DatCat Value</termNote>`

DatCat: **"example"**

**DC identifier:** ISO12620A-0504
**Level(s):** TE LS TS

**Style:** TypedElement
**Vocabulary:** descrip example
**Example:** `<descrip type="example">DatCat Value</descrip>`

DatCat: **"explanation"**

**DC identifier:** ISO12620A-0502
**Level(s):** TE LS TS

**Style:** TypedElement
**Vocabulary:** descrip explanation
**Example:** `<descrip type="explanation">DatCat Value</descrip>`

DatCat: **"external cross reference"**

**DC identifier:** ISO12620A-101807
**Level(s):** TE LS TS

**Style:** TypedElement
**Vocabulary:** xRef externalCrossReference
**Example:** `<xRef type="externalCrossReference">DatCat Value</xRef>`

DatCat: **"false friend"**

**DC identifier:** ISO12620A-0302
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** termNote falseFriend
**Example:** `<termNote type="falseFriend">DatCat Value</termNote>`

DatCat: **"figure"**

**DC identifier:** ISO12620A-050501
**Level(s):** TE LS TS

**Style:** TypedElement
**Vocabulary:** descrip figure
**Example:** `<descrip type="figure">DatCat Value</descrip>`

DatCat: **"foreign text"**

**DC identifier:** ISO12620A-1008
**Level(s):** TE TS

**Style:** Element
**Vocabulary:** foreign
**Example:** `<foreign>DatCat Value</foreign>`

DatCat: **"frequency"**

**DC identifier:** ISO12620A-020304

**Style:** TypedElement

| | |
|---|---|
| **Level(s):** TS | **Vocabulary:** termNote frequency<br>**Example:** `<termNote type="frequency">DatCat Value</termNote>` |

DatCat: "**commonly used**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02030401<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** commonlyUsed<br>**Example:** `commonlyUsed` |

DatCat: "**infrequently used**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02030402<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** infrequentlyUsed<br>**Example:** `infrequentlyUsed` |

DatCat: "**rarely used**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02030403<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** rarelyUsed<br>**Example:** `rarelyUsed` |

DatCat: "**geographical usage**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020302<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote geographicalUsage<br>**Example:** `<termNote type="geographicalUsage">DatCat Value</termNote>` |

DatCat: "**grammatical gender**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020202<br>**Level(s):** TS TCS | **Style:** TypedElement<br>**Vocabulary:** termNote grammaticalGender<br>**Example:** `<termNote type="grammaticalGender">DatCat Value</termNote>` |

DatCat: "**masculine**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02020201<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** masculine<br>**Example:** `masculine` |

DatCat: "**feminine**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02020202<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** feminine<br>**Example:** `feminine` |

DatCat: "**neuter**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02020203<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** neuter<br>**Example:** `neuter` |

DatCat: "**other gender**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02020204<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** otherGender<br>**Example:** `otherGender` |

DatCat: **"grammatical number"**

**DC identifier:** ISO12620A-020203
**Level(s):** TS TCS

**Style:** TypedElement
**Vocabulary:** termNote grammaticalNumber
**Example:** `<termNote type="grammaticalNumber">DatCat Value</termNote>`

DatCat: **"singular"**

**DC identifier:** ISO12620A-02020301
**Level(s):**

**Style:** Simple
**Vocabulary:** singular
**Example:** `singular`

DatCat: **"plural"**

**DC identifier:** ISO12620A-02020302
**Level(s):**

**Style:** Simple
**Vocabulary:** plural
**Example:** `plural`

DatCat: **"dual"**

**DC identifier:** ISO12620A-02020303
**Level(s):**

**Style:** Simple
**Vocabulary:** dual
**Example:** `dual`

DatCat: **"mass noun"**

**DC identifier:** ISO12620A-02020304
**Level(s):**

**Style:** Simple
**Vocabulary:** massNoun
**Example:** `massNoun`

DatCat: **"other number"**

**DC identifier:** ISO12620A-02020305
**Level(s):**

**Style:** Simple
**Vocabulary:** otherNumber
**Example:** `otherNumber`

DatCat: **"grammatical valency"**

**DC identifier:** ISO12620A-020207
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** termNote grammaticalValency
**Example:** `<termNote type="grammaticalValency">DatCat Value</termNote>`

DatCat: **"homograph"**

**DC identifier:** ISO12620A-101805
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** termNote homograph
**Example:** `<termNote type="homograph">DatCat Value</termNote>`

DatCat: **"hotkey"**

**DC identifier:** ISO12620A-100604
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** hi hotkey
**Example:** `<hi type="hotkey">DatCat Value</hi>`

DatCat: **"hyphenation"**

**DC identifier:** ISO12620A-0207

**Style:** TypedElement

| | |
|---|---|
| **Level(s):** TCS | **Vocabulary:** termCompList hyphenation<br>**Example:** `<termCompList type="hyphenation">DatCat Value</termCompList>` |

DatCat: **"index heading"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0905<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** admin indexHeading<br>**Example:** `<admin type="indexHeading">DatCat Value</admin>` |

DatCat: **"keyword"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0904<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** admin keyword<br>**Example:** `<admin type="keyword">DatCat Value</admin>` |

DatCat: **"language identifier"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100701<br>**Level(s):** TE LS TS | **Style:** Attribute<br>**Vocabulary:** xml:lang<br>**Example:** `xml:lang="DatCat Value"` |

DatCat: **"language-planning qualifier"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020902<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote language-planningQualifier<br>**Example:** `<termNote type="language-planningQualifier">DatCat Value</termNote>` |

DatCat: **"recommended term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02090201<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** recommendedTerm<br>**Example:** `recommendedTerm` |

DatCat: **"nonstandardized term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02090202<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** nonstandardizedTerm<br>**Example:** `nonstandardizedTerm` |

DatCat: **"proposed term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02090203<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** proposedTerm<br>**Example:** `proposedTerm` |

DatCat: **"new term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02090204<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** newTerm<br>**Example:** `newTerm` |

DatCat: **"lemma"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020803<br>**Level(s):** TS TCS | **Style:** TypedElement<br>**Vocabulary:** termCompList lemma |

| | |
|---|---|
| | **Example:** `<termCompList type="lemma">DatCat Value</termCompList>` |

DatCat: "**morphological element**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020801<br>**Level(s):** TCS | **Style:** TypedElement<br>**Vocabulary:** termCompList morphologicalElement<br>**Example:** `<termCompList type="morphologicalElement">DatCat Value</termCompList>` |

DatCat: "**normative authorization**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020901<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote normativeAuthorization<br>**Example:** `<termNote type="normativeAuthorization">DatCat Value</termNote>` |

DatCat: "**standardized term**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02090101<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** standardizedTerm<br>**Example:** `standardizedTerm` |

DatCat: "**preferred term**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02090102<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** preferredTerm<br>**Example:** `preferredTerm` |

DatCat: "**admitted term**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02090103<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** admittedTerm<br>**Example:** `admittedTerm` |

DatCat: "**deprecated term**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02090104<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** deprecatedTerm<br>**Example:** `deprecatedTerm` |

DatCat: "**superseded term**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02090105<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** supersededTerm<br>**Example:** `supersededTerm` |

DatCat: "**legal term**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02090106<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** legalTerm<br>**Example:** `legalTerm` |

DatCat: "**regulated term**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02090107<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** regulatedTerm<br>**Example:** `regulatedTerm` |

DatCat: "**note**"

**DC identifier:** ISO12620A-08
**Level(s):** TE LS TS

**Style:** Element
**Vocabulary:** note
**Example:** `<note>DatCat Value</note>`

DatCat: "**originating database**"

**DC identifier:**
**Level(s):**

DatCat: "**originating institution**"

**DC identifier:** ISO12620A-102202
**Level(s):** TE

**Style:** TypedElement
**Vocabulary:** admin originatingInstitution
**Example:** `<admin type="originatingInstitution">DatCat Value</admin>`

DatCat: "**originating person**"

**DC identifier:** ISO12620A-102201
**Level(s):** TE

**Style:** TypedElement
**Vocabulary:** admin originatingPerson
**Example:** `<admin type="originatingPerson">DatCat Value</admin>`

DatCat: "**other binary data**"

**DC identifier:** ISO12620A-050505
**Level(s):** TE LS TS

**Style:** TypedElement
**Vocabulary:** descrip otherBinaryData
**Example:** `<descrip type="otherBinaryData">DatCat Value</descrip>`

DatCat: "**part of speech**"

**DC identifier:** ISO12620A-020201
**Level(s):** TS TCS

**Style:** TypedElement
**Vocabulary:** termNote partOfSpeech
**Example:** `<termNote type="partOfSpeech">DatCat Value</termNote>`

DatCat: "**process status**"

**DC identifier:** ISO12620A-020904
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** termNote processStatus
**Example:** `<termNote type="processStatus">DatCat Value</termNote>`

DatCat: "**unprocessed**"

**DC identifier:** ISO12620A-02090401
**Level(s):**

**Style:** Simple
**Vocabulary:** unprocessed
**Example:** `unprocessed`

DatCat: "**provisionally processed**"

**DC identifier:** ISO12620A-02090402
**Level(s):**

**Style:** Simple
**Vocabulary:** provisionallyProcessed
**Example:** `provisionallyProcessed`

DatCat: "**finalized**"

**DC identifier:** ISO12620A-02090403
**Level(s):**

**Style:** Simple
**Vocabulary:** finalized

| | |
|---|---|
| | **Example:** `finalized` |

DatCat: "**product subset**"

| | |
|---|---|
| **DC identifier:** ISO12620A-100305<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** admin productSubset<br>**Example:** `<admin type="productSubset">DatCat Value</admin>` |

DatCat: "**project subset**"

| | |
|---|---|
| **DC identifier:** ISO12620A-100303<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** admin projectSubset<br>**Example:** `<admin type="projectSubset">DatCat Value</admin>` |

DatCat: "**pronunciation**"

| | |
|---|---|
| **DC identifier:** ISO12620A-0205<br>**Level(s):** TS TCS | **Style:** TypedElement<br>**Vocabulary:** termNote pronunciation<br>**Example:** `<termNote type="pronunciation">DatCat Value</termNote>` |

DatCat: "**proprietary restriction**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020307<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote proprietaryRestriction<br>**Example:** `<termNote type="proprietaryRestriction">DatCat Value</termNote>` |

DatCat: "**trademark**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02030701<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** trademark<br>**Example:** `trademark` |

DatCat: "**service mark**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02030702<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** serviceMark<br>**Example:** `serviceMark` |

DatCat: "**trade name**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02030703<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** tradeName<br>**Example:** `tradeName` |

DatCat: "**quantity**"

| | |
|---|---|
| **DC identifier:** ISO12620A-050701<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** descrip quantity<br>**Example:** `<descrip type="quantity">DatCat Value</descrip>` |

DatCat: "**range**"

| | |
|---|---|
| **DC identifier:** ISO12620A-0507<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** descrip range<br>**Example:** `<descrip type="range">DatCat Value</descrip>` |

DatCat: **"register"**

**DC identifier:** ISO12620A-020303
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** termNote register
**Example:** `<termNote type="register">DatCat Value</termNote>`

DatCat: **"neutral register"**

**DC identifier:** ISO12620A-02030301
**Level(s):**

**Style:** Simple
**Vocabulary:** neutralRegister
**Example:** `neutralRegister`

DatCat: **"technical register"**

**DC identifier:** ISO12620A-02030302
**Level(s):**

**Style:** Simple
**Vocabulary:** technicalRegister
**Example:** `technicalRegister`

DatCat: **"in-house register"**

**DC identifier:** ISO12620A-02030303
**Level(s):**

**Style:** Simple
**Vocabulary:** in-houseRegister
**Example:** `in-houseRegister`

DatCat: **"bench-level register"**

**DC identifier:** ISO12620A-02030304
**Level(s):**

**Style:** Simple
**Vocabulary:** bench-levelRegister
**Example:** `bench-levelRegister`

DatCat: **"slang register"**

**DC identifier:** ISO12620A-02030305
**Level(s):**

**Style:** Simple
**Vocabulary:** slangRegister
**Example:** `slangRegister`

DatCat: **"vulgar register"**

**DC identifier:** ISO12620A-02030306
**Level(s):**

**Style:** Simple
**Vocabulary:** vulgarRegister
**Example:** `vulgarRegister`

DatCat: **"related concept"**

**DC identifier:** ISO12620A-070205
**Level(s):** TE LS

**Style:** TypedElement
**Vocabulary:** descrip relatedConcept
**Example:** `<descrip type="relatedConcept">DatCat Value</descrip>`

DatCat: **"related concept broader"**

**DC identifier:** ISO12620A-07020501
**Level(s):** TE LS

**Style:** TypedElement
**Vocabulary:** descrip relatedConceptBroader
**Example:** `<descrip type="relatedConceptBroader">DatCat Value</descrip>`

DatCat: **"related concept narrower"**

**DC identifier:** ISO12620A-07020502

**Style:** TypedElement

| | |
|---|---|
| **Level(s):** TE LS | **Vocabulary:** descrip relatedConceptNarrower<br>**Example:** `<descrip type="relatedConceptNarrower">DatCat Value</descrip>` |

DatCat: "**reliability code**"

| | |
|---|---|
| **DC identifier:** ISO12620A-0304<br>**Level(s):** TE LS TS | **Style:** TypedElement<br>**Vocabulary:** descrip reliabilityCode<br>**Example:** `<descrip type="reliabilityCode">DatCat Value</descrip>` |

DatCat: "**1**"

| | |
|---|---|
| **DC identifier:** ISO12620A-030401<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** 1<br>**Example:** 1 |

DatCat: "**2**"

| | |
|---|---|
| **DC identifier:** ISO12620A-030402<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** 2<br>**Example:** 2 |

DatCat: "**3**"

| | |
|---|---|
| **DC identifier:** ISO12620A-030403<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** 3<br>**Example:** 3 |

DatCat: "**4**"

| | |
|---|---|
| **DC identifier:** ISO12620A-030404<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** 4<br>**Example:** 4 |

DatCat: "**5**"

| | |
|---|---|
| **DC identifier:** ISO12620A-030405<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** 5<br>**Example:** 5 |

DatCat: "**6**"

| | |
|---|---|
| **DC identifier:** ISO12620A-030406<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** 6<br>**Example:** 6 |

DatCat: "**7**"

| | |
|---|---|
| **DC identifier:** ISO12620A-030407<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** 7<br>**Example:** 7 |

DatCat: "**8**"

| | |
|---|---|
| **DC identifier:** ISO12620A-030408<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** 8<br>**Example:** 8 |

DatCat: "**9**"

**DC identifier:** ISO12620A-030409
**Level(s):**

**Style:** Simple
**Vocabulary:** 9
**Example:** `9`

DatCat: "**10**"

**DC identifier:** ISO12620A-030410
**Level(s):**

**Style:** Simple
**Vocabulary:** 10
**Example:** `10`

DatCat: "**responsibility**"

**DC identifier:** ISO12620A-100202
**Level(s):** TE LS TS TCS

**Style:** TypedElement
**Vocabulary:** transacNote responsibility
**Example:** `<transacNote type="responsibility">DatCat Value</transacNote>`

DatCat: "**sample sentence**"

**DC identifier:** ISO12620A-050302
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** descrip sampleSentence
**Example:** `<descrip type="sampleSentence">DatCat Value</descrip>`

DatCat: "**search term**"

**DC identifier:** ISO12620A-100603
**Level(s):** TE LS TS

**Style:** TypedElement
**Vocabulary:** admin searchTerm
**Example:** `<admin type="searchTerm">DatCat Value</admin>`

DatCat: "**security subset**"

**DC identifier:** ISO12620A-100309
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** admin securitySubset
**Example:** `<admin type="securitySubset">DatCat Value</admin>`

DatCat: "**public**"

**DC identifier:** ISO12620A-10030901
**Level(s):**

**Style:** Simple
**Vocabulary:** public
**Example:** `public`

DatCat: "**confidential**"

**DC identifier:** ISO12620A-10030902
**Level(s):**

**Style:** Simple
**Vocabulary:** confidential
**Example:** `confidential`

DatCat: "**see**"

**DC identifier:** ISO12620A-101801
**Level(s):** TE LS TS TCS

**Style:** TypedElement
**Vocabulary:** ref see
**Example:** `<ref type="see">DatCat Value</ref>`

DatCat: "**sequentially related concept**"

**DC identifier:** ISO12620A-070206

**Style:** TypedElement

| | |
|---|---|
| **Level(s):** TE LS | **Vocabulary:** descrip sequentiallyRelatedConcept<br>**Example:** `<descrip type="sequentiallyRelatedConcept">DatCat Value</descrip>` |

DatCat: **"short form for"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02013004<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote shortFormFor<br>**Example:** `<termNote type="shortFormFor">DatCat Value</termNote>` |

DatCat: **"sort key"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100602<br>**Level(s):** TE LS TS | **Style:** TypedElement<br>**Vocabulary:** admin sortKey<br>**Example:** `<admin type="sortKey">DatCat Value</admin>` |

DatCat: **"source"**

| | |
|---|---|
| **DC identifier:** ISO12620A-1019<br>**Level(s):** TE LS TS | **Style:** TypedElement<br>**Vocabulary:** admin source<br>**Example:** `<admin type="source">DatCat Value</admin>` |

DatCat: **"source identifier"**

| | |
|---|---|
| **DC identifier:** ISO12620A-1020<br>**Level(s):** TE LS TS | **Style:** TypedElement<br>**Vocabulary:** admin sourceIdentifier<br>**Example:** `<admin type="sourceIdentifier">DatCat Value</admin>` |

DatCat: **"source language"**

| | |
|---|---|
| **DC identifier:** ISO12620A-1023<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** admin sourceLanguage<br>**Example:** `<admin type="sourceLanguage">DatCat Value</admin>` |

DatCat: **"source locale"**

| | |
|---|---|
| **DC identifier:**<br>**Level(s):** | |

DatCat: **"source type"**

| | |
|---|---|
| **DC identifier:** ISO12620A-102001<br>**Level(s):** TE LS TS | **Style:** TypedElement<br>**Vocabulary:** adminNote sourceType<br>**Example:** `<adminNote type="sourceType">DatCat Value</adminNote>` |

DatCat: **"parallel text"**

| | |
|---|---|
| **DC identifier:** ISO12620A-10200101<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** parallelText<br>**Example:** `parallelText` |

DatCat: **"background text"**

| | |
|---|---|
| **DC identifier:** ISO12620A-10200102<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** backgroundText<br>**Example:** `backgroundText` |

DatCat: "**spatially related concept**"

**DC identifier:** ISO12620A-070208
**Level(s):** TE LS

**Style:** TypedElement
**Vocabulary:** descrip spatiallyRelatedConcept
**Example:** `<descrip type="spatiallyRelatedConcept">DatCat Value</descrip>`

DatCat: "**subject field**"

**DC identifier:** ISO12620A-04
**Level(s):** TE LS TS

**Style:** TypedElement
**Vocabulary:** descrip subjectField
**Example:** `<descrip type="subjectField">DatCat Value</descrip>`

DatCat: "**subordinate concept generic**"

**DC identifier:** ISO12620A-07020301
**Level(s):** TE LS

**Style:** TypedElement
**Vocabulary:** descrip subordinateConceptGeneric
**Example:** `<descrip type="subordinateConceptGeneric">DatCat Value</descrip>`

DatCat: "**subordinate concept partitive**"

**DC identifier:** ISO12620A-07020302
**Level(s):** TE LS

**Style:** TypedElement
**Vocabulary:** descrip subordinateConceptPartitive
**Example:** `<descrip type="subordinateConceptPartitive">DatCat Value</descrip>`

DatCat: "**subset owner**"

**DC identifier:** ISO12620A-10020210
**Level(s):** TE

**Style:** TypedElement
**Vocabulary:** admin subsetOwner
**Example:** `<admin type="subsetOwner">DatCat Value</admin>`

DatCat: "**superordinate concept generic**"

**DC identifier:** ISO12620A-07020202
**Level(s):** TE LS

**Style:** TypedElement
**Vocabulary:** descrip superordinateConceptGeneric
**Example:** `<descrip type="superordinateConceptGeneric">DatCat Value</descrip>`

DatCat: "**superordinate concept partitive**"

**DC identifier:** ISO12620A-07020203
**Level(s):** TE LS

**Style:** TypedElement
**Vocabulary:** descrip superordinateConceptPartitive
**Example:** `<descrip type="superordinateConceptPartitive">DatCat Value</descrip>`

DatCat: "**syllabification**"

**DC identifier:** ISO12620A-0206
**Level(s):** TCS

**Style:** TypedElement
**Vocabulary:** termCompList syllabification
**Example:** `<termCompList type="syllabification">DatCat Value</termCompList>`

DatCat: "**table**"

**DC identifier:** ISO12620A-050504
**Level(s):** TE LS TS

**Style:** TypedElement
**Vocabulary:** descrip table
**Example:** `<descrip type="table">DatCat Value</descrip>`

DatCat: **"target language"**

**DC identifier:** ISO12620A-1024
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** admin targetLanguage
**Example:** `<admin type="targetLanguage">DatCat Value</admin>`

DatCat: **"target locale"**

**DC identifier:**
**Level(s):**

DatCat: **"temporal qualifier"**

**DC identifier:** ISO12620A-020305
**Level(s):** TS

**Style:** TypedElement
**Vocabulary:** termNote temporalQualifier
**Example:** `<termNote type="temporalQualifier">DatCat Value</termNote>`

DatCat: **"archaic term"**

**DC identifier:** ISO12620A-02030501
**Level(s):**

**Style:** Simple
**Vocabulary:** archaicTerm
**Example:** `archaicTerm`

DatCat: **"outdated term"**

**DC identifier:** ISO12620A-02030502
**Level(s):**

**Style:** Simple
**Vocabulary:** outdatedTerm
**Example:** `outdatedTerm`

DatCat: **"obsolete term"**

**DC identifier:** ISO12620A-02030503
**Level(s):**

**Style:** Simple
**Vocabulary:** obsoleteTerm
**Example:** `obsoleteTerm`

DatCat: **"temporally related concept"**

**DC identifier:** ISO12620A-070207
**Level(s):** TE LS

**Style:** TypedElement
**Vocabulary:** descrip temporallyRelatedConcept
**Example:** `<descrip type="temporallyRelatedConcept">DatCat Value</descrip>`

DatCat: **"term"**

**DC identifier:** ISO12620A-01
**Level(s):** TS TCS

**Style:** Element
**Vocabulary:** term
**Example:** `<term>DatCat Value</term>`

DatCat: **"term element"**

**DC identifier:** ISO12620A-020802
**Level(s):** TCS

**Style:** TypedElement
**Vocabulary:** termCompList termElement
**Example:** `<termCompList type="termElement">DatCat Value</termCompList>`

DatCat: **"term identifier"**

**DC identifier:** ISO12620A-1026

**Style:** Attribute

| | |
|---|---|
| **Level(s):** TS | **Vocabulary:** id<br>**Example:** `id="DatCat Value"` |

DatCat: **"term provenance"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020401<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote termProvenance<br>**Example:** `<termNote type="termProvenance">DatCat Value</termNote>` |

DatCat: **"transdisciplinary borrowing"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02040101<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** transdisciplinaryBorrowing<br>**Example:** `transdisciplinaryBorrowing` |

DatCat: **"translingual borrowing"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02040102<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** translingualBorrowing<br>**Example:** `translingualBorrowing` |

DatCat: **"loan translation"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02040103<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** loanTranslation<br>**Example:** `loanTranslation` |

DatCat: **"neologism"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02040104<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** neologism<br>**Example:** `neologism` |

DatCat: **"term structure"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020804<br>**Level(s):** TCS | **Style:** TypedElement<br>**Vocabulary:** termNote termStructure<br>**Example:** `<termNote type="termStructure">DatCat Value</termNote>` |

DatCat: **"term type"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0201<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote termType<br>**Example:** `<termNote type="termType">DatCat Value</termNote>` |

DatCat: **"entry term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020101<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** entryTerm<br>**Example:** `entryTerm` |

DatCat: **"synonym"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020102<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** synonym<br>**Example:** `synonym` |

DatCat: **"international scientific term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020104<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** internationalScientificTerm<br>**Example:** `internationalScientificTerm` |

DatCat: **"full form"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020107<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** fullForm<br>**Example:** `fullForm` |

DatCat: **"transcribed form"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020111<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** transcribedForm<br>**Example:** `transcribedForm` |

DatCat: **"symbol"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020113<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** symbol<br>**Example:** `symbol` |

DatCat: **"formula"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020114<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** formula<br>**Example:** `formula` |

DatCat: **"equation"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020115<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** equation<br>**Example:** `equation` |

DatCat: **"logical expression"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020116<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** logicalExpression<br>**Example:** `logicalExpression` |

DatCat: **"common name"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020105<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** commonName<br>**Example:** `commonName` |

DatCat: **"abbreviated form of term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020108<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** abbreviatedFormOfTerm<br>**Example:** `abbreviatedFormOfTerm` |

DatCat: **"variant"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020109 | **Style:** Simple |

| | |
|---|---|
| **Level(s):** | **Vocabulary:** variant<br>**Example:** `variant` |

DatCat: "**short form of term**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02010802<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** shortFormOfTerm<br>**Example:** `shortFormOfTerm` |

DatCat: "**transliterated form**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020110<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** transliteratedForm<br>**Example:** `transliteratedForm` |

DatCat: "**sku**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02011701<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** sku<br>**Example:** `sku` |

DatCat: "**part number**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02011702<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** partNumber<br>**Example:** `partNumber` |

DatCat: "**phraseological unit**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020118<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** phraseologicalUnit<br>**Example:** `phraseologicalUnit` |

DatCat: "**synonymous phrase**"

| | |
|---|---|
| **DC identifier:** ISO12620A-02011803<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** synonymousPhrase<br>**Example:** `synonymousPhrase` |

DatCat: "**standard text**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020119<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** standardText<br>**Example:** `standardText` |

DatCat: "**string**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020120<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** string<br>**Example:** `string` |

DatCat: "**internationalism**"

| | |
|---|---|
| **DC identifier:** ISO12620A-020106<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** internationalism<br>**Example:** `internationalism` |

DatCat: **"terminology management transactions"**

| | |
|---|---|
| **DC identifier:** ISO12620A-1001<br>**Level(s):** TE LS TS TCS | **Style:** Element<br>**Vocabulary:** transac<br>**Example:** `<transac>DatCat Value</transac>` |

DatCat: **"origination"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100101<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** origination<br>**Example:** `origination` |

DatCat: **"input"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100102<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** input<br>**Example:** `input` |

DatCat: **"modification"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100103<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** modification<br>**Example:** `modification` |

DatCat: **"check"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100104<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** check<br>**Example:** `check` |

DatCat: **"approval"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100105<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** approval<br>**Example:** `approval` |

DatCat: **"withdrawal"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100106<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** withdrawal<br>**Example:** `withdrawal` |

DatCat: **"standardization"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100107<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** standardization<br>**Example:** `standardization` |

DatCat: **"exportation"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100108<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** exportation<br>**Example:** `exportation` |

DatCat: **"importation"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100109 | **Style:** Simple |

| | |
|---|---|
| **Level(s):** | **Vocabulary:** importation<br>**Example:** `importation` |

DatCat: **"proposal"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100110<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** proposal<br>**Example:** `proposal` |

DatCat: **"user access"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100111<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** userAccess<br>**Example:** `userAccess` |

DatCat: **"thesaurus descriptor"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0902<br>**Level(s):** TE | **Style:** TypedElement<br>**Vocabulary:** descrip thesaurusDescriptor<br>**Example:** `<descrip type="thesaurusDescriptor">DatCat Value</descrip>` |

DatCat: **"time restriction"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020306<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote timeRestriction<br>**Example:** `<termNote type="timeRestriction">DatCat Value</termNote>` |

DatCat: **"transfer comment"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0305<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote transferComment<br>**Example:** `<termNote type="transferComment">DatCat Value</termNote>` |

DatCat: **"unit"**

| | |
|---|---|
| **DC identifier:** ISO12620A-0506<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** descrip unit<br>**Example:** `<descrip type="unit">DatCat Value</descrip>` |

DatCat: **"usage count"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100203<br>**Level(s):** TE | **Style:** TypedElement<br>**Vocabulary:** transacNote usageCount<br>**Example:** `<transacNote type="usageCount">DatCat Value</transacNote>` |

DatCat: **"usage note"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020301<br>**Level(s):** TS | **Style:** TypedElement<br>**Vocabulary:** termNote usageNote<br>**Example:** `<termNote type="usageNote">DatCat Value</termNote>` |

DatCat: **"video"**

| | |
|---|---|
| **DC identifier:** ISO12620A-050503<br>**Level(s):** TE LS TS | **Style:** TypedElement<br>**Vocabulary:** descrip video<br>**Example:** `<descrip type="video">DatCat Value</descrip>` |

#### B.5 Conclusion

MSC qualifies as a TML in compliance with this International Standard. Furthermore, related formats that are also TMLs can be defined simply by selecting a different set of data categories, without varying the XML expansion trees, styles, and vocabularies. Hence the TMLs thus defined share a common core structure. As a result, MSC forms the basis for the MSC-type subset of all possible TMLs. This subset constitutes a family of related formats, some of which do not involve blind interchange, which can in many cases be processed by a single software tool.

## Annex C (normative): The Geneter Terminology Markup Language

#### C.1 Introduction

This appendix presents the characteristics of a Terminology Markup Language conformant with this standard. The complete model (DTD and schemas) and examples of its usage can be found on the Maintenance Agency and Registration Authority site.

The Geneter format is one possible Terminology Markup Language (TML) that can be realized using the Terminology Markup Framework (TMF) as defined by ISO 16642.

Geneter provides an XML-based implementation of the "meta model" as defined by this standard for Terminological Data Collections (TDC).

Geneter is a format that describes data categories and their relationships in a TDC. It is "generic" as far as it takes into account data categories that have been defined in ISO 12620 (Data Categories) or identified in real applications.

More restricted subsets corresponding to the needs of specific groups of users can be derived from this format. These subsets can be "closed" so that they can allow for "blind interchange".

A "Geneter subset" **shall** comply with the "subsetting rules" as defined in C.6.1

To be "Geneter conformant" a Geneter instance **shall** conform to the Geneter format or to a Geneter subset. A conformant instance **shall** be a valid XML document as defined in http://www.w3.org/TR/2000/REC-xml-20001006 ("An XML document is valid if it has an associated document type declaration and if the document complies with the constraints expressed in it.")

This annex presents an example illustrating the conformance of the Geneter TML to the Terminology Markup Framework (C.1) and then specifies the Geneter implementation of each meta-model level: Global Information Section (C.3), Terminological Entry (C.4), Complementary Information (C.5). The last part (C.6) deals with methods for restricting or extending the Geneter format.

#### C.2 Example : specification of a Geneter subset as a TML

The last part of this annex (C.6.4) specifies a Geneter subset corresponding to a specific data structure related to user needs. The resulting model, an encoded example and the vocabulary and style specifications of this TML are illustrated below.

##### C.2.1 Geneter subset

```
<!-- structure definition --><!ELEMENT geneter (terminologicalEntry+) ><!ATTLIST
geneter profile (oracle) 'oracle'><!ELEMENT terminologicalEntry (SubjectField,
Definition, languageCtn+) ><!ATTLIST terminologicalEntry

     identifier CDATA #IMPLIED>

<!ELEMENT languageCtn (Term*)><!ATTLIST languageCtn value (en|hu) #IMPLIED>
<!-- data categories definition --><!ELEMENT SubjectField (#PCDATA)> <!ELEMENT
Definition (#PCDATA)> <!ATTLIST Definition xml:lang (en) 'en' ><!ELEMENT Term (#PCDATA)
><!ATTLIST Term formType (fullForm|abbreviation) #IMPLIED

     workingStatus (working|consolidated) #REQUIRED>
```

##### C.2.2 XML encoding of an entry

```
<?xml version ="1.0" encoding="ISO-8859-1" ?><!DOCTYPE geneter SYSTEM 'oracle.dtd' []>
<geneterprofile = "oracle">

     <terminologicalEntry identifier = 'ID67' workingStatus='consolidated'>
          <SubjectField>Manufacturing</SubjectField>
          <Definition xml:lang ='en'>A value beetween 0 and 1 used in ...</Definition>
          <languageCtn value='en'>
               <Term formType='fullForm'>alpha smoothing</Term>
          </languageCtn>
          <languageCtn value='hu'>
               <Term> Alfa simítási tényezõ</Term>
          </languageCtn>
     </terminologicalEntry>

</geneter>
```

#### C.2.3 Geneter vocabulary and style

| **12620** | **Style** | **Geneter** |
|---|---|---|
| entry type | | terminologicalEntry |
| entry identifier | **Level** : TE<br>**Type** : attribut<br>**Anchor information** : terminologicalEntry<br>**Value** : txt | terminologicalEntry identifier |
| subjectfield | **Level** : TE<br>**Type** : element<br>**Value** : txt | SubjectField |
| definition | **Level** : TE<br>**Type** : element<br>**Value** : txt | Definition |
| ? | **Level** : TE<br>**Type** : attribute<br>**Anchor information** : Definition<br>**Value** : txt | xml:lang |
| ? | | languageCtn |
| language identifier | **Level** : LS<br>**Type** : attribute<br>**Anchor information** : languageCtn<br>**Value** : txt | languageCtn value |
| term | **Level** : TS<br>**Type** : element<br>**Value** : txt | Term |
| term related information | **Level** : TS<br>**Type** : attribute<br>**Anchor information** : Term<br>**Value** : picklist | Term formType |
| element working status | **Level** : TS<br>**Type** : attribute<br>**Anchor information** : Term<br>**Value** : picklist | Term workingStatus |

#### C.3 Global Information Section (GIS)

The GIS of a Geneter TDC contains information (meta data) about the collection. It is represented by a `<header>` element. Meta data are defined by two name spaces according to the Dublin Core qualified meta data model (http://dublincore.org/documents/2000/07/11/dcmes-qualifiers )

```
xmlns:dc="http://purl.oclc.org/dc#" xmlns:dcq="http://purl.org/dc/qualifiers/1.0/"
```

**Name spaces declaration for a Geneter header**

```
<header>

     <meta name = "DC.Type" scheme = "DCMIType" content = "Dataset">

     <meta name = "DC.Date.Issued" scheme = "ANSI.X3.X30-1985" content = "20011511">

</header>
```

**Example of a Geneter header**

#### C.4 Terminological entry (TE)

A terminological entry consists of data categories and containers for further data categories.

##### C.4.1 Data category types

The Geneter format is based on three types of data categories:

- 1) **Structural** data categories (for instance: `<Definition>`, `<Term>` or `<PartOfSpeech>`) used to describe the terminological information. They are listed with their name, attributes and content model in table C.4.9, column 2.
- 2) **Embedded** data categories used within the content of a structural data category to incorporate terminology related information Their content model is `%Line;` (see C.4.5.1 for mixed content model types).

| **name** | **attributes** | **explanation** |
|---|---|---|
| Annotation | **type** : CDATA<br>**scheme** : CDATA<br>**value** : CDATA<br>%GeneralAttributes; | Additional linguistic information |
| Characteristic | **type** : CDATA<br>%GeneralAttributes; | A.5.8 |
| EntailedTerm | **type** : CDATA<br>%GeneralAttributes; | A.10.6.1 |
| ForeignText | **type** : CDATA<br>**language** : CDATA<br>%GeneralAttributes; | A.10.8 |
| Keyword | %GeneralAttributes; | A.9.4.2 |

**Embedded elements**

- 3) **Basic** (non specific) elements used within the content of a structural data category or of an embedded data category. (The `<Ptr>` element is described in C.4.6).

| **name** | **attributes** | **content model** | **explanation** |
|---|---|---|---|
| Data | %generalAttributes | #PCDATA | simple data category for basic text |
| Date | **type :** CDATA<br>**scheme :** CDATA<br>**calendar :** CDATA<br>%GeneralAttributes; | CDATA | A.10.2.1 |
| Fpi | **type** : CDATA<br>%GeneralAttributes; | %URI; | A.10.21.2 |
| Quantity | **type** : CDATA<br>**value :** CDATA<br>**unit :** CDATA<br>%GeneralAttributes; | EMPTY | |
| Range | **type** : CDATA<br>**min** : CDATA<br>**max** : CDATA<br>%GeneralAttributes; | EMPTY | A.5.7 |
| Segment | **type** : CDATA<br>%GeneralAttributes; | <seg>+ (CDATA) | Segmented element (hyphenation, syllabification...) |
| Unit | **type** : CDATA | %basicLine; | A.5.6 |

| | | | |
|---|---|---|---|
| | %GeneralAttributes; | | |
| Url | **type** : CDATA<br>%GeneralAttributes; | %URI | A.10.21.1 |
| Where | **type** : CDATA<br>%GeneralAttributes; | %basicLine; | A location indicator |
| Who | **type** : CDATA<br>**role** : CDATA<br>%GeneralAttributes; | %basicLine; | Person or corporate body involved in some action |

**Basic elements**

```
<Definition>A liquid with a <Characteristic>boiling temperature</Characteristic> of
<Quantity value="100", Unit = "Celsius"/></Definition>
```

**Example with the three types of data category**

#### C.4.2 Characteristics of a Geneter data category: name, attributes, content model

For each data category, the Geneter format specifies its **name** (an XML Generic Identifier), its **attributes** and its **content model**.

For practical reasons, structural data categories are grouped into blocks corresponding respectively to

- data **function** (administrative information, description of a property, description of a relation)
- data **position** in the tree structure of a terminological entry (corresponding to the four levels TE, LS, TS, TCS of the meta model).

For instance, at the TE level (Language Independent Section), elements are grouped in the `%lisAdminDatCat; %lisPropDatCat; %lisRelDatCat;` blocks. The prefix for the other levels are: "lds" (Language Section), "term" (Term Section) and "component" (Term Component Section).

#### C.4.3 General attributes

The attributes for any data category are:

- the core attributes (id, class, style, title) and the internationalization attributes (xml:lang, dir) defined in XHTML1 (http://www.w3.org/TR/xhtml1 ).
- specific attributes such as "security" or "workingStatus" corresponding respectively to ISO 12620 "element Working Status" (10.11) and "security subset" (10.3.9)

#### C.4.4 Container

A container is a structure used to "**refine**" a data category (i.e. to supply additional information). A container begins with the data category to be refined. This data category is unique and compulsory. The name of a container is formed by adding the suffix "Ctn"to the name of this data category. Example:

```
<termCtn>

     <Term>barbed wire</Term>

     <Source>source</Source>

</termCtn>
```

**Example of refinement by the <Source>**

```
<registerCtn value='neuter'>

     <Note>...</Note>
```

```
</registerCtn>
```

**Example of container for an empty element**

Containers can be embedded :

```
<termCtn>

     <Term>barbed wire</Term>

     <sourceCtn>
          <Source>source</Source>
          <Note>note on source</Note>
     </sourceCtn>

</termCtn>
```

**Example of embedded containers**

A data category can be recursive (i.e. self-refining) :

```
<subjectFieldCtn>

     <SubjectField>chemistry</SubjectField>

     <subjectFieldCtn>
          <SubjectField>petrol</SubjectField>
     <subjectFieldCtn>

<subjectFieldCtn>
```

#### C.4.5 Content models

The content model of an element can be expressed as either mixed content or composite content. These models are described below.

##### C.4.5.1 mixed content

There are four kinds of mixed content respectively named `%Flow;, %Inline;, %Line;, %basicLine;.`

The first two are defined as an extension of XHTML 1.0. Using Modularization for XHTML (http://www.w3.org/TR/xhtml-modularization ), the `%inline;` entity of the HTML "Text Module" is redefined to incorporate embedded data categories (`%embeddedDataCategories;` and basic data categories (`%basicDataCategories;`). So that they can occur anywhere the HTML `%inline` entity; is used.

```
<!ENTITY % inline "a | %special; | %fontstyle; | %phrase; | %inline.forms;">
```

**Original XHTML %inline declaration**

```
<!ENTITY % inline "a | %special; | %fontstyle; | %phrase;| Data | Ptr | %
embeddedDataCategories; | %basicDataCategories;">
```

**Modified XHTML %inline declaration**

The original `%inline;` entity augmented with all the basic elements is renamed as `%Line;`. It is used as a content model for embedded elements.

The original %inline; entity is renamed as %basicLine. It is used as a content model for basic elements.

```
<Example>

     boiling points

     <table><tr><td>
          <Characteristic value= "100" unit = "Celsius" />
     </td></tr></table>

</Example>
```

**Example of enriched mixed content (with an embedded data category in a table)**

#### C.4.5.2 composite content

Composite content is created using an ordered set of enumerated elements. These elements are either structural data categories (in the case of recursivity for instance) or basic data categories.

```
<Importation>

     <Date>01-02-2001</Date>

     <Where>EUROTERMS</Where>

</Importation>
```

**Example of a composite element**

#### C.4.6 Pointers

Relations between a terminological entry and Complementary Information (see C.5) are expressed by the <ptr> element. This element conforms to the xlink specification (http://www.w3.org/TR/xlink).

```
<Owner>

     <ptr xml:lang = 'en' xlink:embed = 'none' xlink:href = 'person01'> BROWN, J.</ptr>

</Owner>
```

**Example of a pointer using the <ptr> element**

#### C.4.7 Structure of a terminological entry

The structure of a Geneter terminological entry consists of three embedded containers (languageCtn, termCtn and termComponantCtn):

| Meta model | Geneter |
|---|---|
| TerminologicalEntry (TE) | terminologicalEntry |
| LanguageSection (LS) | languageCtn |
| TermSection (TS) | termCtn |
| TermComponantSection (TCS) | termComponantCtn |

**Correspondence between TMF Anchor levels and Geneter**

#### C.4.8 Example of terminological entry

```
<terminologicalEntry identifier='07'>

     <Owner><Ptr xlink:href = 'person01'>xxx</Ptr></Owner>

     <languageCtn value ='en'>
          <Definition>wire with short, sharp points on it</Definition>
          <termCtn>
               <Term>barbed wire</Term>
               <componentCtn
                    <Word>barbed</Word>
                    <PartOfSpeech>adj</PartOfSpeech>
               </componentCtn>
          </termCtn>
     </languageCtn>

</terminologicalEntry>
```

**An instance**

| Geneter encoding | Explanation | Level |
|---|---|---|
| <terminologicalEntry identifier='TE07'> | beginning of the terminological entry with an attribute for the identifier | TE |
| <Owner><Ptr xlink:href = 'person01'>xxx</Ptr></Owner> | data category <Owner> with a link towards the description of a person | |
| <languageCtn value ='en'> | container for a <Language Section> with a language attribute | LS |
| <Definition>wire with short, sharp points on it</Definition> | data category <Definition>, content = "wire with... " | |
| <termCtn> | container for the description of a term and its complements | TS |
| <Term>barbed wire</Term> | data category <Term>, content = "barbed wire" | |
| <componentCtn | container for a Component group with an attribute "rank" indicating the position of the component inside the term | TCS |
| <Word>barbed</Word> | data category <Word> content = "barbed" | |
| <PartOfSpeech>adj</PartOfSpeech> | data category <Grammar>, content = "adj" | |
| </componentCtn> | end of the component container | |
| </termCtn> | end of the term container | |
| </terminologicalEntry> | end of the terminological entry | |

**Detailed explanation**

#### C.4.9 The tree structure of a terminological entry

##### C.4.9.1. Geneter synopsis

The following tree represents the Geneter name, attributes, content model, position of any data category in the Geneter structure as well as the ISO 12620 position from which it is derived. Non ISO 12620 elements and entities (name given to repetitive information) are also defined below.

| Num | Element name | Attribut | Content model | ISO 12620 |
|---|---|---|---|---|
| 1 | terminologicalEntry | **identifier:** %URI<br>**type**: %terminological EntryType;<br>**GeneterVersion:** CDATA<br>**profile:** %URI;<br>**profileVersion:** CDATA | (%lis; \| %lds;)* | A.10.10 |

| | | | | |
|---|---|---|---|---|
| | | %GeneralAttributes; | | |
| 1.1. | %lis;<br>(Language Independent Section) | %GeneralAttributes; | (%lisAdminDatCat; \| %lisPropDatCat; \| %lisRelDatCat;)* | |
| 1.1.1. | %lisAdminDatCat; | | | |
| 1.1.1.1 | Application | %GeneralAttributes; | %Inline; | A.10.3.6 |
| | applicationCtn | %GeneralAttributes; | Application, (Application \| %adminAgent; \| %cpt;)* | |
| 1.1.1.2 | Approval | **status:** CDATA<br>%GeneralAttributes; | %act; | A.10.2.1.5,<br>A.10.2.2.5 |
| | approvalCtn | %GeneralAttributes; | Approval, (%cpt;)* | |
| 1.1.1.3 | BusinessUnit | %GeneralAttributes; | %Inline; | A.10.3.8 |
| | businessUnitCtn | %GeneralAttributes; | BusinessUnit, (BusinessUnit \| %adminItem; \| %cpt;)* | |
| 1.1.1.4 | Check | **status:** CDATA<br>%GeneralAttributes; | %act; | A.10.2.1.4,<br>A.10.2.2.4 |
| | checkCtn | %GeneralAttributes; | Check, (%cpt;)* | |
| 1.1.1.5 | Contributor | **role:** %contributorRole; | %Inline; | (1) |
| | contributorCtn | %GeneralAttributes | Contributor, (%Cpt;)* | |
| 1.1.1.6 | Copyright | %GeneralAttributes; | %Inline; | B.13 |
| | copyrightCtn | %GeneralAttributes; | Copyright, (%cpt;)* | |
| 1.1.1.7 | Coverage | %GeneralAttributes; | %Inline; | (2) |
| | coverageCtn | %GeneralAttributes; | Coverage, (%Cpt;)* | |
| 1.1.1.8 | Customer | **type:** CDATA<br>%GeneralAttributes; | %Inline; | A.10.3.1 |
| | customerCtn | %GeneralAttributes; | Customer, (Customer \| %adminItem; \| %cpt;)* | |
| 1.1.1.9 | EntrySource | %GeneralAttributes; | %Inline; | A.10.13 |
| | entrySourceCtn | %GeneralAttributes; | EntrySource, (%cpt;)* | |
| 1.1.1.10 | Environment | %GeneralAttributes; | %Inline; | A.10.3.7 |
| | environmentCtn | %GeneralAttributes; | Environment, (Environment \| %adminAgent; \| %cpt;)* | |
| 1.1.1.11 | Exportation | %GeneralAttributes; | %act;,where* | A.10.2.1.8,<br>A.10.2.2.8 |
| | exportationCtn | %GeneralAttributes; | exportation, (%cpt;)* | |
| 1.1.1.12 | FileIdentifier | %GeneralAttributes; | %Inline; | A.10.17 |
| | fileIdentifierCtn | %GeneralAttributes; | FileIdentifier, (%cpt;)* | |
| 1.1.1.13 | Input | %GeneralAttributes; | (%act;)* | A.10.2.1.2,<br>A.10.2.2.2 |
| | inputCtn | %GeneralAttributes; | Input, (%cpt;)* | |
| 1.1.1.14 | Importation | %GeneralAttributes; | %act;,where? | A.10.2.1.9,<br>A.10.2.2.9 |
| | importationCtn | %GeneralAttributes; | Importation, (%cpt;)* | |
| 1.1.1.15 | LastModification | %GeneralAttributes; | (%act;)* | (3) |
| | lastModificationCtn | %GeneralAttributes; | LastModification, (%cpt;)* | |
| 1.1.1.16 | Modification | %GeneralAttributes; | (%act;)* | A.10.2.1.3,<br>A.10.2.2.3 |
| | modificationCtn | %GeneralAttributes; | Modification, (%cpt;)* | |
| 1.1.1.17 | Note | **type:** %NoteType ;<br>%GeneralAttributes; | %Flow; | A.8 |
| | noteCtn | %GeneralAttributes; | Note, (Source\|sourceCtn)* | |
| 1.1.1.18 | Origination | %GeneralAttributes; | (%act;)* | A.10.2.1.1, |

| | | | | |
|---|---|---|---|---|
| | | | | A.10.2.2.1 |
| | originationCtn | %GeneralAttributes; | Origination, (%cpt;)* | |
| 1.1.1.19 | Owner | %GeneralAttributes; | %Inline; | 10.2.2.10 |
| | OwnerCtn | %GeneralAttributes; | Owner, (%adminItem%cpt;)* | |
| 1.1.1. 20 | Product | %GeneralAttributes; | %Inline; | A.10.3.5 |
| | ProductCtn | %GeneralAttributes; | Product, (Product \| % adminAgent; \| %cpt;)* | |
| 1.1.1.21 | Project | **type:** CDATA %GeneralAttributes; | %Inline; | A.10.3.3 |
| | ProjectCtn | %GeneralAttributes; | Project, (Project \| % adminAgent; \| %cpt;)* | |
| 1.1.1.22 | RecordIdentifier | %GeneralAttributes; | %Inline; | A.10.16 |
| | recordIdentifierCtn | %GeneralAttributes; | RecordIdentifier, (%cpt;)* | |
| 1.1.1.23 | ReliabilityCode | **scale:** CDATA %GeneralAttributes; | %Inline; | A.3.4 |
| | reliabilityCodeCtn | %GeneralAttributes; | ReliabilityCode, (%cpt;)* | |
| 1.1.1. 24 | Responsibility | **type:** %ResponsibilityType; %GeneralAttributes; | (%act;)* | A.10.2.2 |
| | responsibilityCtn | %GeneralAttributes; | Responsibility, (%cpt;)* | |
| 1.1.1.25 | Source | **type:** CDATA %GeneralAttributes; | %Inline; | A.10.19 |
| | sourceCtn | %GeneralAttributes; | Source, (Note\|noteCtn)* | |
| 1.1.1.26 | SourceLanguage | **source:** CDATA %GeneralAttributes; | EMPTY | (4) |
| | sourceLanguageCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.1.1.27 | StandardizationStatus | **status:** CDATA %GeneralAttributes; | %act; | A.10.2.1.7 |
| | standardizationStatusCtn | %GeneralAttributes; | Standardization, (%cpt;)* | |
| 1.1.1.28 | TargetLanguage | **target:** CDATA %GeneralAttributes; | EMPTY | (5) |
| | targetLanguageCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.1.1.29 | User | %GeneralAttributes; | %Inline; | A.10.2.2.6 |
| | userCtn | %GeneralAttributes; | User, (%cpt;)* | |
| 1.1.1.30 | Withdrawal | %GeneralAttributes; | (%act;)* | A.10.2.1.6, A.10.2.2.7 |
| | withdrawalCtn | %GeneralAttributes; | Withdrawal, (%cpt;)* | |
| 1.1.2 | %lisPropDatCat; | | | |
| 1.1.2.1 | ConceptPosition | **position:** CDATA %GeneralAttributes; | %Inline; | A.7.2 |
| | conceptPositionCtn | %GeneralAttributes; | ConceptPosition, (%cpt;*) | |
| 1.1.2.2 | DefiningCharacteristic | **type:** CDATA %GeneralAttributes; | %Inline; | A.5.8 |
| | definingCharacteristicCtn | %GeneralAttributes; | DefiningCharacteristic, (% cpt;*) | |
| 1.1.2.3 | Definition | **type:** %DefinitionType; %GeneralAttributes; | %Flow; | A.5.1 |
| | definitionCtn | %GeneralAttributes; | Definition, ((%cpt;)* \| ParallelText \| parallelTextCtn) * | |
| 1.1.2.4 | Example | %GeneralAttributes; | %Flow; | A.5.4 |
| | exampleCtn | %GeneralAttributes; | Example, ((%cpt;)* \| | |

| | | | | |
|---|---|---|---|---|
| | | | ParallelText \| parallelTextCtn)* | |
| 1.1.2.5 | Explanation | %GeneralAttributes; | %Flow; | A.5.2 |
| | explanationCtn | %GeneralAttributes; | Explanation, ((%cpt;)* \| ParallelText \| parallelTextCtn)* | |
| 1.1.2.6 | InternationalScientificTerm | %GeneralAttributes; | %Inline; | A.2.1.4 |
| | internationalScientificTermCtn | %GeneralAttributes; | InternationalScientificTerm, (RelatedDescription \| %cpt;)* | |
| 1.1.2.7 | Illustration | **type:** %IllustrationType ; **media**:%IllustrationMediaType ; %GeneralAttributes; | %Flow; | A.5.5 |
| | illustrationCtn | %GeneralAttributes; | Illustration, (%cpt;)* | |
| 1.1.2.8 | Scope | %GeneralAttributes; | %Inline; | (6) |
| | scopeCtn | %GeneralAttributes; | Scope, (%cpt;)* | |
| 1.1.2.9 | SubjectField | **type:** %SubjectFieldType ; **codedValue**: CDATA %GeneralAttributes; | %Inline; | A.4 |
| | subjectFieldCtn | %GeneralAttributes; | (SubjectField, ((%cpt;)* \| SubjectField \| subjectFieldCtn)*) | |
| 1.1.3 | %lisRelDatCat; | | | |
| 1.1.3.1 | Antonym | **type:** %AntonymType ; %GeneralAttributes; | %Inline; | A.10.18.6 |
| | antonymCtn | %GeneralAttributes; | Antonym, (%cpt;)* | |
| 1.1.3.2 | AssociativeRelation | **type:** CDATA %GeneralAttributes; | %Inline; | A.6.4 |
| | associativeRelationCtn | %GeneralAttributes; | AssociativeRelation, (%cpt;)* | |
| 1.1.3.3 | CausalRelation | **type:** % CausalRelationType; %GeneralAttributes; | %Inline; | (7) |
| | causalRelationCtn | %GeneralAttributes; | CausalRelation, (%cpt;)* | |
| 1.1.3.4 | DegreeOfEquivalence | **value:** %DegreeOf EquivalenceValue; **sourceLanguage:** CDATA **targetLanguage:** CDATA %GeneralAttributes; | %Inline; | A.3.1 |
| | degreeOfEquivalenceCtn | %GeneralAttributes; | DegreeOfEquivalence, (% cpt;)* | |
| 1.1.3.5 | GenericRelation | **type:** % GenericRelationType %GeneralAttributes; | %Inline; | A.6.1 |
| | genericRelationCtn | %GeneralAttributes; | GenericRelation, (%cpt;)* | |
| 1.1.3.6 | PartitiveRelation | **type:** % PartitiveRelationType; %GeneralAttributes; | %Inline; | A.6.2 |
| | partitiveRelationCtn | %GeneralAttributes; | PartitiveRelation, (%cpt;)* | |
| 1.1.3.7 | SpatialRelation | **type:** % SpatialRelationType; %GeneralAttributes; | %Inline; | A.6.3.2 |
| | spatialRelationCtn | %GeneralAttributes; | SpatialRelation, (%cpt;)* | |
| 1.1.3.8 | RelatedDescription | **type:** %Related DescriptionType; %GeneralAttributes; | %Inline; | (8) |
| | relatedDescriptionCtn | %GeneralAttributes; | RelatedDescription, (%cpt;) | |
| | | | | |

| | | | | |
|---|---|---|---|---|
| 1.1.3.8 | TemporalRelation | **type:** %Temporal RelationType; %GeneralAttributes; | %Inline; | A.6.3.1 |
| | temporalRelationCtn | %GeneralAttributes; | TemporalRelation, (%cpt;)* | |
| | %Ids;<br>(Language Dependent Section) | | (Language \| languageCtn) | |
| 1.2 | languageCtn | **value:** CDATA<br>**identifier:** %URI;<br>%GeneralAttributes; | (ExternalLanguageSectionCtn \| (%ldsAdminDatCat; \| % ldsPropDatCat; \| % ldsRelDatCat; \| %free; \| % term;)* | (11) |
| 1.2.0 | externalLanguageSectionCtn | **value:** CDATA<br>%GeneralAttributes; | (%cpt;)* | |
| 1.2.0.1 | ExternalLanguageSection | **value:** CDATA<br>**location:** CDATA<br>**identifier:** %URI; | EMPTY | (12) |
| 1.2.1 | %ldsAdminDatCat; | | | |
| 1.2.1.1 - 1.2.1.30 | Same as 1.1.1.1 to 1.1.1.30 | | | |
| 1.2.2 | %ldsPropDatCat; | | | |
| 1.2.2.1 | ConceptPosition | **position:** CDATA<br>%GeneralAttributes; | %Inline; | A.7.2 |
| | conceptPositionCtn | %GeneralAttributes; | ConceptPosition, (%cpt;*) | |
| 1.2.2.2 | Context | **type:** %ContextType ;<br>%GeneralAttributes; | %Flow; | A.5.3 |
| | contextCtn | %GeneralAttributes; | Context, ((%cpt;)* \| ParallelText \| parallelTextCtn)* | |
| 1.2.2.3 | DefiningCharacteristic | **type:** CDATA<br>%GeneralAttributes; | %Inline; | A.5.8 |
| | definingCharacteristicCtn | %GeneralAttributes; | DefiningCharacteristic, (% cpt;*) | |
| 1.2.2.4 | Definition | **type:** %DefinitionType ;<br>%GeneralAttributes; | %Flow; | A.5.1 |
| | definitionCtn | %GeneralAttributes; | Definition, ((%cpt;)* \| ParallelText \| parallelTextCtn)* | |
| 1.2.2.5 | Example | %GeneralAttributes; | %Flow; | A.5.4 |
| | exampleCtn | %GeneralAttributes; | Example, ((%cpt;)* \| ParallelText \| parallelTextCtn)* | |
| 1.2.2.6 | Explanation | %GeneralAttributes; | %Flow; | A.5.2 |
| | explanationCtn | %GeneralAttributes; | Explanation, ((%cpt;)* \| ParallelText \| parallelTextCtn)* | |
| 1.2.2.7 | Illustration | **type:** %IllustrationType ;<br>**media**:% IllustrationMediaType ;<br>%GeneralAttributes; | %Flow; | A.5.5 |
| | illustrationCtn | %GeneralAttributes; | Illustration, (%cpt;)* | |
| 1.2.2.8 | Scope | %GeneralAttributes; | %Inline; | (6) |
| | scopeCtn | %GeneralAttributes; | Scope, (%cpt;)* | |
| 1.2.2.9 | SearchTerm | %GeneralAttributes; | %Inline; | A.10.6.3 |
| | searchTermCtn | %GeneralAttributes; | SearchTerm, (%cpt;)* | |
| 1.2.2.10 | SortKey | %GeneralAttributes; | %Inline; | A.10.6.2 |
| | | | | |

| | | | | |
|---|---|---|---|---|
| | sortKeyCtn | %GeneralAttributes; | SortKey, (%cpt;)* | |
| 1.2.2.11 | SubjectField | **type:** %SubjectField Type ; **codedValue**: CDATA %GeneralAttributes; | %Inline; | A.4 |
| | subjectFieldCtn | %GeneralAttributes; | (SubjectField, ((%cpt;)* \| SubjectField \| subjectFieldCtn)*) | |
| 1.2.3 | %ldsRelDatCat; | | | |
| 1.2.3.1 | Antonym | **type:** %AntonymType; %GeneralAttributes; | %Inline; | A.10.18.6 |
| | antonymCtn | %GeneralAttributes; | Antonym, (%cpt;)* | |
| 1.2.3.2 | AssociativeRelation | **type:** CDATA %GeneralAttributes; | %Inline; | A.6.4 |
| | associativeRelationCtn | %GeneralAttributes; | AssociativeRelation, (%cpt;)* | |
| 1.2.3.3 | CausalRelation | **type:** % CausalRelationType; %GeneralAttributes; | %Inline; | (7) |
| | causalRelationCtn | %GeneralAttributes; | CausalRelation, (%cpt;)* | |
| 1.2.3.4 | DegreeOfEquivalence | **value:** %DegreeOf EquivalenceValue; **targetLanguage**: CDATA %GeneralAttributes; | EMPTY | A.3.1 |
| | degreeOfEquivalenceCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.3.5 | GenericRelation | **type:** % GenericRelationType; %GeneralAttributes; | %Inline; | A.6.1 |
| | genericRelationCtn | %GeneralAttributes; | GenericRelation, (%cpt;)* | |
| 1.2.3.6 | PartitiveRelation | **type:** % PartitiveRelationType; %GeneralAttributes; | %Inline; | A.6.2 |
| | partitiveRelationCtn | %GeneralAttributes; | PartitiveRelation, (%cpt;)* | |
| 1.2.3.7 | RelatedDescription | **type:** % RelatedDescriptionType; %GeneralAttributes; | %Inline; | (8) |
| | relatedDescriptionCtn | %GeneralAttributes; | RelatedDescription, (%cpt;) | |
| 1.2.3.8 | SpatialRelation | **type:** % SpatialRelationValue; %GeneralAttributes; | %Inline; | A.6.3.2 |
| | spatialRelationCtn | %GeneralAttributes; | SpatialRelation, (%cpt;)* | |
| 1.2.3.9 | TemporalRelation | **type:** % TemporalRelationType; %GeneralAttributes; | %Inline; | A.6.3.1 |
| | temporalRelationCtn | %GeneralAttributes; | TemporalRelation, (%cpt;)* | |
| | | | | |
| 1.2.5. | %term; | | Term \| termCtn | |
| 1.2.5.1 | termCtn | %GeneralAttributes; | (Term, (%termAdminDatCat; \| %termPropDatCat; \| % termRelDatCat; \| Term \| termCtn \| %free; \| % component;)* | |
| 1.2.5.2 | Term | **identifier:** %URI; **designationType**: % TermDesignationType; **type**: %TermType; **status**: %TermStatus; **transcript:**%transcript; | %Inline; | A.1 |

| | | | | |
|---|---|---|---|---|
| | | **layOut**: %TermLayout; **formType**: %TermForm Type; **variantType**: %Term VariantType; **degreeOfSynonymy**: % TermDegreeOfSynomy; %GeneralAttributes; | | |
| 1.2.5.3. | %termAdminDatCat; | | | |
| 1.2.5.3.1 - 1.2.5.3.32 | same as %ldsAdminDatCat; | | | |
| 1.2.5.4. | %termPropDatCat; | | | |
| 1.2.5.4.1 | Animacy | **value:** %AnimacyValue; %GeneralAttributes; | EMPTY | A.2.2.4 |
| | animacyCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.2 | Collocation | %GeneralAttributes; | %Inline; | A.2.1.18.1 |
| | collocationCtn | %GeneralAttributes; | Collocation, (%cpt;)* | |
| 1.2.5.4.3 | Context | **type:** %ContextType; %GeneralAttributes; | %Flow; | A.5.3 |
| | contextCtn | %GeneralAttributes; | Context, ((%cpt;)* \| ParallelText \| parallelTextCtn) * | |
| 1.2.5.4.4 | Definition | **type:** %DefinitionType; %GeneralAttributes; | %Flow; | A.5.1 |
| | definitionCtn | %GeneralAttributes; | Definition, ((%cpt;)* \| ParallelText \| parallelTextCtn) * | |
| 1.2.5.4.5 | Derivation | **type:** %DerivationType; %GeneralAttributes; | %Inline; | (13) |
| | derivationCtn | %GeneralAttributes; | Derivation, (%cpt;)* | |
| 1.2.5.4.6 | Etymology | **language:** CDATA %GeneralAttributes; | %Inline; | A.2.4.2 |
| | etymologyCtn | %GeneralAttributes; | Etymology, (%cpt;)* | |
| 1.2.5.4.7 | Example | %GeneralAttributes; | %Flow; | A.5.4 |
| | exampleCtn | %GeneralAttributes; | Example, (%cpt;)* | |
| 1.2.5.4.8 | Explanation | %GeneralAttributes; | %Flow; | A.5.2 |
| | explanationCtn | %GeneralAttributes; | Explanation, (%cpt;)* | |
| 1.2.5.4.9 | Frequency | **value:** %FrequencyValue; %GeneralAttributes; | EMPTY | A.2.3.4 |
| | frequencyCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.10 | GeographicalUsage | **value:** CDATA **usage**:% GeographicalUsageType; %GeneralAttributes; | EMPTY | A.2.3.2 |
| | geographicalUsageGr | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.11 | Grammar | **type:** CDATA %GeneralAttributes; | %Inline; | A.2.2 |
| | grammarCtn | %GeneralAttributes; | Grammar, (%cpt;)* | |
| 1.2.5.4.12 | GrammaticalGender | **value:** %Grammatical GenderValue; %GeneralAttributes; | EMPTY | A.2.2.2 |
| | grammaticalGenderCtn | **value:** CDATA | (%cpt;)* | |

| | | | | |
|---|---|---|---|---|
| | | %GeneralAttributes; | | |
| 1.2.5.4.13 | GrammaticalNumber | **value:** %Grammatical NumberValue; %GeneralAttributes; | EMPTY | A.2.2.3 |
| | grammaticalNumberCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.14 | Hyphenation | %GeneralAttributes; | %Inline; | A.2.7 |
| | hyphenationCtn | %GeneralAttributes; | Hyphenation, (%cpt;)* | |
| 1.2.5.4.15 | Illustration | **type:** %IllustrationType **media**: %Illustration Media; %GeneralAttributes; | | A.5.5 |
| | illustrationCtn | %GeneralAttributes; | Illustration, (%cpt;)* | |
| 1.2.5.4.16 | Inflection | **type:** %InflectionType; %GeneralAttributes; | %Inline; | (14) |
| | inflectionCtn | %GeneralAttributes; | Inflection, (%cpt;)* | |
| 1.2.5.4.17 | InHouseRegister | **value:** %RegisterValue; %GeneralAttributes; | %Inline; | A.2.3.3 |
| | inHouseRegisterCtn | %GeneralAttributes; | InHouse, (%cpt;)* | |
| 1.2.5.4.18 | LanguagePlanningQualifier | **value:** %Language PlanningQualifierValue; %GeneralAttributes; | EMPTY | A.2.9.2 |
| | languagePlanningQualifierCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.19 | NormativeAuthorization | **value:** %Normative AuthorizationValue; %GeneralAttributes; | EMPTY | A.2.9.1 |
| | normativeAuthorizationCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.20 | PartOfSpeech | **value:** CDATA %GeneralAttributes; | EMPTY | A.2.2.1 |
| | partOfSpeechCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.21 | Pronunciation | %GeneralAttributes; | %Inline; | A.2.5 |
| | pronunciationCtn | %GeneralAttributes; | Pronunciation, (%cpt;)* | |
| 1.2.5.4.22 | ProprietaryRestriction | **value:** %Proprietary RestrictionValue; %GeneralAttributes; | EMPTY | A.2.3.7 |
| | proprietaryRestrictionCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.23 | Register | **value:** %RegisterValue; %GeneralAttributes; | EMPTY | A.2.3.3 |
| | registerCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.24 | Scope | %GeneralAttributes; | %Inline; | (6) |
| | scopeCtn | %GeneralAttributes; | Scope, (%cpt;)* | |
| 1.2.5.4.25 | SearchTerm | %GeneralAttributes; | %Inline; | A.10.6.3 |
| | searchTermCtn | %GeneralAttributes; | SearchTerm, (%cpt;)* | |
| 1.2.5.4.26 | SortKey | %GeneralAttributes; | %Inline; | A.10.6.2 |
| | sortKeyCtn | %GeneralAttributes; | SortKey, (%cpt;)* | |
| 1.2.5.4.27 | SubjectField | **type:** %SubjectFieldType; **codedValue**: CDATA %GeneralAttributes; | %Inline; | A.4 |
| | subjectFieldCtn | %GeneralAttributes; | (SubjectField, ((%cpt;)* \| SubjectField \| subjectFieldCtn )*) | |

| | | | | |
|---|---|---|---|---|
| 1.2.5.4.28 | Syllabification | %GeneralAttributes; | %Inline; | A.2.6 |
| | syllabificationCtn | %GeneralAttributes; | Syllabification, (%cpt;)* | |
| 1.2.5.4.29 | SyntacticalFunction | %GeneralAttributes; | %Inline; | (15). |
| | syntacticalFunctionCtn | %GeneralAttributes; | SyntacticalFunction, (%cpt;)* | |
| 1.2.5.4.30 | TemporalQualifier | **value:** %Temporal QualifierValue; %GeneralAttributes; | EMPTY | A.2.3.5 |
| | temporalQualifierCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.4.31 | TermComplement | **type:** %Term ComplementType; **transcript:**%transcript; %GeneralAttributes; | (%Inline;)* | (16) |
| | termComplementCtn | | TermComplement, (%cpt;)* | |
| 1.2.5.4.32 | TermDisplay | **transcript:** %transcript; % GeneralAttributes; | %Inline; | (19) |
| | termDisplayCtn | %GeneralAttributes; | TermDisplay, (%cpt;)* | |
| 1.2.5.4.33 | TermProvenance | **type:** %Term ProvenanceType ; **language** CDATA %GeneralAttributes; | %Inline; | A.2.4.1 |
| | termProvenanceCtn | %GeneralAttributes; | TermProvenance, (%cpt;)* | |
| 1.2.5.4.34 | TimeRestriction | %GeneralAttributes; | %Inline; | A.2.3.6 |
| | timeRestrictionCtn | %GeneralAttributes; | TimeRestriction, (%cpt;)* | |
| 1.2.5.5. | %termRelDatCat; | | | |
| 1.2.5.5.1 | Antonym | **type:** %AntonymType; %GeneralAttributes; | %Inline; | A.10.18.6 |
| | antonymCtn | %GeneralAttributes; | Antonym, (%cpt;)* | |
| 1.2.5.5.2 | AssociativeRelation | **type:** CDATA %GeneralAttributes; | %Inline; | A.6.4 |
| | associativeRelationCtn | %GeneralAttributes; | AssociativeRelation, (%cpt;)* | |
| 1.2.5.5.3 | CausalRelation | **type:** % CausalRelationType; %GeneralAttributes; | %Inline; | (7) |
| | causalRelationCtn | %GeneralAttributes; | CausalRelation, (%cpt;)* | |
| 1.2.5.5.4 | DegreeOfEquivalence | **value:** %DegreeOf EquivalenceValue; **directionality**: %Degree OfEquivalenceDirectionality; **targetLanguage**: CDATA %GeneralAttributes; | EMPTY | A.3.1 |
| | degreeOfEquivalenceCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.5.5 | DegreeOfSynonymy | **value:** %DegreeO fSynonymyValue; %GeneralAttributes; | %Inline; | A.2.10 |
| | degreeOfSynonymyCtn | %GeneralAttributes; | DegreeOfSynonymy, (%cpt;)* | |
| 1.2.5.5.6 | FalseFriend | **targetLanguage:** CDATA %GeneralAttributes; | %Inline; | A.3.2 |
| | falseFriendCtn | %GeneralAttributes; | FalseFriend, (%cpt;)* | |
| 1.2.5.5.7 | FormOfTerm | **type:** %FormOf TermType; %GeneralAttributes; | %Inline; | A.2.1.7, A.2.1.8 |
| | formOfTermCtn | %GeneralAttributes; | FormOfTerm, (%cpt;)* | |
| 1.2.5.5.8 | GenericRelation | **type:** %generic | %Inline; | A.6.1 |

| | | | | |
|---|---|---|---|---|
| | | RelationType; %GeneralAttributes; | | |
| | genericRelationCtn | %GeneralAttributes; | GenericRelation, (%cpt;)* | |
| 1.2.5.5.9 | Homograph | **indicator:** CDATA %GeneralAttributes; | %Inline; | A.10.18.5 |
| | homographCtn | %GeneralAttributes; | Homograph, (%cpt;)* | |
| 1.2.5.5.10 | Homonym | **indicator:** CDATA %GeneralAttributes; | %Inline; | (18) |
| | HomonymCtn | %GeneralAttributes; | Homonym, (%cpt;)* | |
| 1.2.5.5.11 | Homophone | %GeneralAttributes; | %Inline; | (19) |
| | HomophoneCtn | %GeneralAttributes; | Homophone, (%cpt;)* | |
| 1.2.5.5.12 | PartitiveRelation | **type:** % PartitiveRelationType; %GeneralAttributes; | %Inline; | A.6.2 |
| | partitiveRelationCtn | %GeneralAttributes; | PartitiveRelation, (%cpt;)* | |
| 1.2.5.5.13 | Polysemy | %GeneralAttributes; | %Inline; | (20) |
| | polysemyCtn | %GeneralAttributes; | Polysemy, (%cpt;)* | |
| 1.2.5.5.14 | RelatedDescription | **type:** % RelatedDescriptionType; **targetLanguage**: CDATA %GeneralAttributes; | %Inline; | (8) |
| | relatedDescriptionCtn | %GeneralAttributes; | RelatedDescription, (%cpt;) | |
| 1.2.5.5.15 | SpatialRelation | **type:** %Spatial RelationType; %GeneralAttributes; | %Inline; | A.6.3.2 |
| | spatialRelationCtn | %GeneralAttributes; | SpatialRelation, (%cpt;)* | |
| 1.2.5.5.16 | TemporalRelation | **type:** %Temporal RelationType; %GeneralAttributes; | %Inline; | A.6.3.1 |
| | temporalRelationCtn | %GeneralAttributes; | TemporalRelation, (%cpt;)* | |
| 1.2.5.5.16 | Variant | **type:** %TermVariantType; **directionality**: % VariantDirectionality; %GeneralAttributes; | %Inline; | A.2.1.9 |
| | variantCtn | %GeneralAttributes; | Variant, (%cpt;)* | |
| | %component; | | Component \| componentCtn | |
| 1.2.5.7. | componentCtn | identifier: %URI; %GeneralAttributes; | (Component, (%componentAdminDatCat; \| %componentPropDatCat; \| % componentRelDatCat; \| % free;)* | |
| 1.2.5.7.1 | Component | **rank:** CDATA **type**: CDATA **transcript:**%transcript; %GeneralAttributes; | %Inline; | A.2.8.2 |
| 1.2.5.7.2 | %componentAdminDatCat; | | | |
| 1.2.5.7.2.1 - 1.2.5.7.2.32 | Same as %termAdminDatCat; | | | |
| 1.2.5.7.3 | %componentPropDatCat; | | | |
| 1.2.5.7.3.1 | Animacy | **value:** %AnimacyValue; %GeneralAttributes; | EMPTY | A.2.2.4 |
| | animacyCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.7.3.2 | Derivation | **type:** %DerivationType; %GeneralAttributes; | %Inline; | (13) |

| | | | | |
|---|---|---|---|---|
| | derivationCtn | %GeneralAttributes; | Derivation, (%cpt;)* | |
| 1.2.5.7.3.3 | Etymology | **targetLanguage:** CDATA %GeneralAttributes; | %Inline; | A.2.4.2 |
| | etymologyCtn | %GeneralAttributes; | Etymology, (%cpt;)* | |
| 1.2.5.7.3.4 | Grammar | **type:** CDATA %GeneralAttributes; | %Inline; | A.2.2 |
| | grammarCtn | %GeneralAttributes; | Grammar, (%cpt;)* | |
| 1.2.5.7.3.5 | GrammaticalGender | **value:** %Grammatical GenderValue; %GeneralAttributes; | EMPTY | A.2.2.2 |
| | grammaticalGenderCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.7.3.6 | GrammaticalNumber | **value:** %Grammatical NumberValue; %GeneralAttributes; | EMPTY | A.2.2.3 |
| | grammaticalNumberCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.7.3.7 | Hyphenation | %GeneralAttributes; | %Inline; | A.2.7 |
| | hyphenationCtn | %GeneralAttributes; | Hyphenation, (%cpt;)* | |
| 1.2.5.7.3.8 | Inflection | **type:** %InflectionType; %GeneralAttributes; | %Inline; | (14) |
| | inflectionCtn | %GeneralAttributes; | Inflection, (%cpt;)* | |
| 1.2.5.7.3.9 | PartOfSpeech | **value:** CDATA %GeneralAttributes; | EMPTY | A.2.2.1 |
| | partOfSpeechCtn | **value:** CDATA %GeneralAttributes; | (%cpt;)* | |
| 1.2.5.7.3.10 | Pronunciation | %GeneralAttributes; | %Inline; | A.2.5 |
| | pronunciationCtn | %GeneralAttributes; | Pronunciation, (%cpt;)* | |
| 1.2.5.7.3.11 | SearchTerm | %GeneralAttributes; | %Inline; | A.10.6.3 |
| | searchTermCtn | %GeneralAttributes; | SearchTerm, (%cpt;)* | |
| 1.2.5.7.3.12 | SortKey | %GeneralAttributes; | %Inline; | A.10.6.2 |
| | sortKeyCtn | %GeneralAttributes; | SortKey, (%cpt;)* | |
| 1.2.5.7.3.13 | Syllabification | %GeneralAttributes; | %Inline; | A.2.6 |
| | syllabificationCtn | %GeneralAttributes; | Syllabification, (%cpt;)* | |
| 1.2.5.7.3.14 | SyntacticalFunction | %GeneralAttributes; | %Inline; | (15) |
| | syntacticalFunctionCtn | %GeneralAttributes; | SyntacticalFunction, (%cpt;)* | |
| 1.2.5.7.4 | %componentRelDatCat; | | | |
| 1.2.5.7.4.1 | FalseFriend | **targetLanguage:** CDATA %GeneralAttributes; | %Inline; | A.3.2 |
| | falseFriendCtn | %GeneralAttributes; | FalseFriend, (%cpt;)* | |
| 1.2.5.7.4.2 | FormOfTerm | **type:** %FormOfTermType; %GeneralAttributes; | %Inline; | A.2.1.7, A.2.1.8 |
| | formOfTermCtn | %GeneralAttributes; | FormOfTerm, (%cpt;)* | |
| 1.2.5.7.4.3 | Homograph | **indicator:** CDATA %GeneralAttributes; | %Inline; | A.10.18.5 |
| | homographCtn | %GeneralAttributes; | Homograph, (%cpt;)* | |
| 1.2.5.7.4.4 | Homonym | **indicator:** CDATA %GeneralAttributes; | %Inline; | (18) |
| | homonymCtn | %GeneralAttributes; | Homonym, (%cpt;)* | |
| 1.2.5.7.4.5 | Homophone | %GeneralAttributes; | %Inline; | (19) |
| | homophoneCtn | %GeneralAttributes; | Homophone, (%cpt;)* | |
| 1.2.5.7.4.6 | Polysemy | %GeneralAttributes; | %Inline; | (20) |

| | | | | |
|---|---|---|---|---|
| | polysemyCtn | %GeneralAttributes; | Polysemy, (%cpt;)* | |
| 1.2.5.7.4.7 | RelatedDescription | **type:** %Related DescriptionType;<br>**targetLanguage**: CDATA<br>%GeneralAttributes; | %Inline; | (8) |
| | relatedDescriptionCtn | %GeneralAttributes; | RelatedDescription, (%cpt;) | |

#### C.4.9.2 Non 12620 data categories

(1) **Contributor** = Any person or organization having a role in the production of the item.
(2) **Coverage** = The extent or scope of the content of the resource. Coverage will typically include spatial location.
(3) **LastModification** = Responsibility and date of the last modification of data.
(4) **SourceLanguage** = In an entry, the language in which a concept has been designated originally.
(5) **TargetLanguage** =. In an entry, the languages in which equivalent designations are provided.
(6) **Scope** = Further indications about the field of application of a concept.
(7) **CausalRelation** = Associative relation between a cause and its effect. [ISO 1087].
(8) **RelatedDescription** = Link with a non terminological description of a term (dictionary, lexicon) or a concept (thesaurus, ontology).
(9) **Free** = see C.6.5
(10) **FreeVal** = see C.6.5
(11) **languageCtn** = A container describing a concept in one language.
(12) **externalLanguageSection** = A language container located on an remote device.
(13) **Derivation** = process of new word formation through the Modification (addition, deletion or remplacement) of a morpheme (suffix) a stem (root).
(14) **Inflection** = Modification of a word with elements that express some grammatical aspects and relations.
(15) **SyntacticalFunction** = Function of a term or a word in the relationships between linguistic units or in the grammatical construction.
(16) **TermComplement** = Ancillary part of a term (the "to" preposition for an English verb for instance).
(17) **TermDisplay** = A displayable or printable form of a term (including embedded grammatical information for instance).
(18) **Homonym** = Terms having an identical pronunciation and/or spelling but referring to different concepts.
(19) **Homophone** = Terms having an identical pronunciation but different spellings and referring to different concepts.
(20) **Polysemy** = Characteristic of a sign (un signifier) that has several contents, several values and several meanings.

#### C.4.9.3 Entities for content models

*(square brackets contain a ISO 12620 position or a short explanation)*

**%act;**: 'date?, who*' *[date and responsibility of a transaction]*
**%adminAgent;**: 'BusinessUnit | businessUnitCtn|Contributor | contributorCtn | Customer | customerCtn|Owner | ownerCtn' *[administrative information about person or organizations]*
**%adminItem;**:
'Project|projectCtn|Product|productCtn|Application|applicationCtn|Environment|environmentCtn' *[administrative information about applications]*
**%cpt ;**: (Note|noteCtn|Source|sourceCtn)* *[complements to a data category inside a container]*
**%free;** : (Free | freeCtn)* *[used for negotiated interchange of extra data categories]*
**%URI** : *[content type for a Uniform Resource Locator (http://www.w3.org/TR/uri-clarification)]*

#### C.4.9.4 Entities for suggested picklists

**%AnimacyValue**: 'animate | inanimate' *[A.2.2.4]*
**%AntonymType**: 'antonymComplement | antonymContrast' *[A.10.18.6]*
**%CausalRelationType**: 'cause | consequence' *[causal relation between concepts]*
**%ComplementType**: 'ante | pos' *[ancillary part of a term]*
**%ContextType**: 'definingContext | explanatoryContext | associativeContext | linguisticContext | metalinguisticContext' *[A.5.3]*
**%ContributorRole**: 'expert | proposer' *[role of contributor to a work]*
**%DefinitionType**: 'intensionalDefinition | extensionalDefinition | partitiveDefinition' *[A.5.1]*
**%DegreeOfEquivalenceDirectionality**: 'bidirectional | monodirectional' *[A.3.3]*
**%DegreeOfEquivalenceValue**: 'narrower | equivalent | quasiEquivalent | broader | nonEquivalent' *[A.3.1]*
**%DegreeOfSynonymyValue**: 'narrower | synonymous | quasiSynonymous | broader | nonsynonymous' *[A.2.10]*
**%DerivationType**: 'regressive | learned | improper'
**%FormOfTermType**: 'fullForm | abbreviation | shortFormOfTerm | initialism | acronym | clippedTerm' *[A.2.1.7 A.2.1.8]*
**%FrequencyValue**: 'commonly | infrequently | rarely' *[A.2.3.4]*
**%GenericRelationType**: 'superordinateConcept | subordinateConcept | coordinateConcept' *[A.6.1]*
**%GeographicalUsageType**: 'used | nonUsed' *[A.3.2]*
**%GrammaticalGenderValue**: 'masculine | feminine | neuter' *[A.2.2]*
**%GrammaticalNumberValue**: 'singular | plural | dual | massNoun' *[A.2.3]*

**%IllustrationMediaType**: 'image | audio | video' *[A.5.5]*
**%IllustrationType**: 'symbol | formula | equation | logicalExpression | figure' *[2.1.13 - 2.1.16]*
**%InflectionType**: 'root | verbal | nominal | pronominal' *[type of modification of a word]*
**%LanguagePlanningQualifie**r: 'recommendedTerm | nonstandardizedTerm | proposedTerm | newTerm' *[A.2.9.2]*
**%NormativeAuthorizationValue**: 'standardizedTerm | preferredTerm | admittedTerm | deprecatedTerm | prohibitedTerm | superseded Term | legalTerm | regulatedTerm' *[A.2.9.1]*
**%NoteType**: 'linguisticNote | technicalNote | userNote | workingNote | transferComment' *[A.8]*
**%PartitiveRelationType**: 'broaderConcept | narrowerConcept' *[A.6.2]*
**%ProcessStatusValue**: 'unprocessed | provisionallyProcessed | finalized' *[A.2.9.4]*
**%ProprietaryRestrictionValue**: 'trademark | tradeName' *[A.2.3.7]*
**%RegisterValue**: 'neutral | technical | benchLevel | slang | vulgar | familiar' *[A.2.3.3]*
**%RelatedDescriptionList**: 'ontology | thesaurus | documentaryLanguage |dictionary | lexicon | translationMemoryData' *[non terminological description of terms or concepts]*
**%ResponsibilityType**: 'person|corporateBody'
**%SpatialRelationType**: 'backward | forward | contiguous' *[spatial relation between concepts]*
**%SubjectFieldType**: 'classificationNumber | indexHeading' *[A.4]*
**%TemporalQualifierValue**: 'archaicTerm | outdatedTerm | obsoleteTerm' *[A.2.3.5]*
**%TemporalRelationType**: 'Preceding | Succeeding | Coincident' *[temporal relation between concepts]*
**%TermDesignationType**: 'term | formula | symbol | equation | logicalExpression' *[2.1.13 - 2.1.16]*
**%TermDegreeOfSynonymy**: 'narrower | broader' *[A.2.10]*
**%TermFormType**: 'fullForm | abbreviation | shortFormOfTerm | initialism | acronym | clippedTerm' *[A.2.1.7 A.2.1.8]*
**%terminologicalEntryType**: 'conceptEntry | standardizedEntry | collocation | phrase | setPhrase | standardText | synonymousPhrase | neologism | geographicalName | commonName | properName | collectiveName | officialDenomination | parallelSegment | managementUnit | partNumber' *[A.10.10]*
**%TermLayout**: 'main | secondary' *[administrative status of a term]*
**%TermProvenanceType**: 'transdisciplinaryBorrowing | translingualBorrowing | loanTranslation | shiftInMeaning' *[A.2.4.1]*
**%TermStatus**: ' neologism | wordCreation | foreignDesignation ' *[status of a new term]*
**%TermType**: 'collocation | formula | phrase | setPhrase | standardText | synonymousPhrase | internationalism | internationalScientificTerm | geographicalName | commonName | properName | collectiveName | officialDenomination | managementUnit | partNumber' *[A.2]*
**%TermVariantType**: 'orthographical|grammatical' *[A.2.1.9]*
**%TransScript**: 'transcribedForm|transliterated|romanizedForm' *[A.2.1.10 - A.2.1.12]*
**%VariantDirectionality**: 'isVariantOf|hasForVariant' *[directionality for variants]*

#### C.5 Complementary Information (CI)

##### C.5.1 Geneter Complementary Information types

- bibliographical information based on ISO 690 (`<monographEntry>`, `<partOfMonographEntry>`, `<contributionEntry>`, `<serialEntry>`, `<articleEntry>`, `<patentEntry>`, `<standardEntry>`, `<audioVideoEntry>`)
- bibliographical information based on ISO 690-2 for electronic documents (`<eMonographEntry>`, `electronic docments: <ePartOfMonographEntry>`,`<eContributionEntry>`, `<eSerialEntry>`, `<ArticleEntry>`, `<eElectronicBoardEntry>`, `<eMessageEntry>`)
- bibliographical information based on ISO 12083 (`<refDocEntry>`)
- description of persons, (`<person>`) based on ISO 12083 bibliographic description;
- description of corporate bodies, (`<corporateBody>`) based on ISO 12083 bibliographic description;
- description of thesaurus based on ISO 2788-1986 "Guidelines for the establishment and development of monolingual thesauri" and ISO 5964:1985 "Guidelines for the establishment and development of multilingual thesauri" (`<thesaurus>`);
- description of machine readable dictionaries based on ISO 1951:1997.
- description of ontologies based on a specialization of the Ontology Inference Layer (http://www.ontoknowledge.org/oil );
- XHTML documents (`<xymlPage>`);
- transitory language containers for exchanging information about one concept in one language (`<transitoryLanguageCtn>`);
- collating sequences [A.10.9]
- encoded binary data for exchanging image or sound or any other non-XML document (doc, pdf, html...), (`<encodedFile>`);
- other objects, (`<freeObject>`).

##### C.5.2 A mechanism for extending Complementary Information

By using XML name spaces, other types of linguistic description can be included in a Geneter collection. This mechanism can be used to manage lexicons (OLIF format: http://www.olif.net ), parallel segments for machine translation (TMX: http://www.lisa.org/tmx , XLIFF: http://www.xliff.org ) and specialized ontologies (OIL: http://www.ontoknowledge.org/oil ). The `<RelatedDescription>` element links terminological entries with these descriptions.

#### C.6 Geneter restriction and extension

For particular needs, it is possible to create subsets based on the Geneter format. Any instance of a Geneter subset **must** be valid against the Geneter DTD. A Geneter subset **must** have a required "profile" attribute giving the Uniform Resource Locator of the subset model. To be compatible with the general model, a subset **must** comply with the following general rules of XML:

##### C.6.1 Rules to constitute a subset :

*For data elements :*

- 1 - any element which has an occurrence indicator ? or * can be deleted ;
- 2 - any element occurrence indicator (?, *, +) can be deleted ;
- 3 - when two elements are combined by OR connector " | " in a content model, one of the two elements can be deleted ;
- 4 - the occurrence indicator * can be replaced by the occurrence indicators ? or +.

*For attributes :*

- 5 - the attributes whose default value is not the key word #REQUIRED can be deleted ;
- 6 - when the attribute value comes from an enumerated list, the list can be reduced but it must contain at least one value ;
- 7 - when the attribute value is cdata type, cdata can be replaced by an enumerated list.

##### C.6.2 Different types of subset :

- a subset which contains neither element nor attribute "free" is a "strict" subset of Geneter.
- if a subset contains <Free> elements whose type and value are literal or taken from enumerated lists, the subset is "closed". Such a subsetcould be called a "jargon" of Geneter.
- if a subset contains <Free> elements whose type and value is CDATA, the subset is "open".

##### C.6.3 Blind subset

By applying rule 1 of subsetting to fuzzy data categories like <Grammar> and by applying it to all the <Free> data categories and to the <free> content element, it is possible to design a more concise Geneter model for blind interchange purpose. The subset mentioned in C.6.2 is such a blind subset.

##### C.6.4 Building a subset : an example

This example is based on a flat source structure in which all the "fields" of data are delimited by a comma. In order to create a Geneter subset corresponding to the original structure and a Geneter instance of these data, the four steps are:

- to identify the type of each element
- to map each of them to a Geneter position
- to design a Geneter subset able to host (and only) all these positions
- to encode the sample

###### C.6.4.1 A data sample

```
67, Manufacturing,,Standard,alpha smoothing factor,Approved,A value between 0 and 1
used in statistical forecasting calculations for smoothing demand fluctuations. ORACLE
Inventory uses the factor to determine how much weight to give to current demand when
calculating a forecast.,Alfa simitási tényezõ
```

###### C.6.4.2 1st step : Analysis of the elements

| **Data category** | **Data** | **12620 correspondence** |
|---|---|---|
| EntryNumber | 67 | entry identifier (A.10.15) |
| Domain | Manufacturing | Subject field (A.4) |
| Product | | Product subset (A.10.3.3) |

| | | |
|---|---|---|
| Datatype (a full form as opposed to an abbreviation) | Standard | Termtype (A.2.1) |
| English | alpha smoothing factor | Term (A.1) |
| Status (an indication of the administrative status of the Hungarian term) | Approved | Process status (A.2.9.4) |
| Definition | A value between 0 and 1 used in statistical forecasting calculations for smoothing demand fluctuations. | Definition (A.5.1) |
| Hungarian term | Alfa simitási tényezõ | Term (A.1) |

#### C.6.4.3 2nd step : The projection scheme from the source format to Geneter

| Data category | Geneter equivalent |
|---|---|
| EntryNumber | 1 terminologicalEntry (identifier attribute) |
| Domain | 1.1.2.7 terminologicalEntry/SubjectField |
| Product | 1.1.1.22 terminologicalEntry/Product |
| Datatype | 1.2.5.2 terminologicalEntry/languageCtn/Term (formType attribute) |
| English term | 1.2.5.2 terminologicalEntry/languageCtn/Term |
| Status of the Hungarian term | 1.2.5.2 terminologicalEntry/languageCtn/Term (workingStatus attribute |
| Definition | 1.1.2.1 terminologicalEntry/Definition |
| Hungarian term | 1.2.5.2 terminologicalEntry/languageCtn/Term |

**Note:** Definition has been put in the Language Independent Section because it applies to the whole entry. Datatype is a property, not a relation, so it is encoded as an attribute (formType).

For the two following steps (designing and encoding a subset), see C.2.

#### C.6.5 Geneter extensions and negotiated interchange

For specific needs, new data categories can be added to the Geneter model at each level of the structure or inside the content models. If XML validity is required for an interchange transaction, these elements **must** be transformed into the meta data category `<Free>` or into a container `<freeCtn>` which are defined in the Geneter format. The negotiation process consists of exchanging the semantics of these free elements with the partner receiving the data.

For instance, a structural data category for indicating the unit rate for a data item can be defined in the Geneter model by the statement:

```
<!ELEMENT Rate (Quantity)>
```

**Definition of an extension**

This element has to be added at some level of the Geneter tree (the `%lisAdminDatCat;` block for instance because it is an administrative information characterizing the whole entry). A possible instance of this element will be:

```
<Rate><Quantity value = "5" unit = "US Dollar"/></Rate>
```

**Encoding of an extension for local management**

For exchange purposes this extra element will be transformed as follows (by an XSLT style-sheet for instance). This encoding is conformant to the Geneter definition of a `<Free>` element. It will validate against the Geneter model.

```
<Free type = "Rate"><Quantity value = "5" unit = "US Dollar"/></Free>
```

**Encoding of an extension for negotiated interchange**

## Annex D (Informative): Conformance of terminological data to TMF

This annex discusses how XML-based terminological data can be made conformant to TMF by analysing the structure and content of the the data and performing certain transformations of these data. The end result of this analysis is the specification of a TML that both represents the terminological data without loss of information, and is interoperable with other TMLs as specified in this standard.

### D.1 Sample terminological data

Consider the following example XML-based representation of a terminological entry from an Automotive Engineering terminology database:

```
<termBank>
     <tbid>00aa</tbid>
     <tbDescription>Automotive Engineering</tbDescription>
     <conceptEntry>
          <domainOfConcept>ABS</domainOfConcept>
          <conceptLastModified>21-08-2001</conceptLastModified>
          <termGroup>
               <languageCode>Deutsch</languageCode>
               <termDefinition> Bauteile, die die elektronischen Steuer- und
               Regelvorgänge für die Blockierregelung und die
               Antriebsschlupfregelung übernehmen.</termDefinition>
               <termString>ABS/ASR-Steuerung</termString>
               <usageDescriptors>
                    <usedIn>Germany</usedIn>
                    <usedIn>Switzerland</usedIn>
               </usageDescriptors>
               <wordClass>n</wordClass>
               <wordGender>f</wordGender>
               <termLastModified>21-08-2001</termLastModified>
          </termGroup>
          <termGroup>
               <languageCode>English</languageCode>
               <termString>ABS/ASR control</termString>
               <usageDescriptors>
                    <usedIn>Britain</usedIn>
               </usageDescriptors>
               <wordClass>n</wordClass>
               <termLastModified>20-08-2001</termLastModified>
          </termGroup>
     </conceptEntry>
</termBank>
```

### D.2 Description of content of elements

The following table describes the information contained within this example:

| XML Element | Description | Description of contents |
|---|---|---|
| <tbid> | Unique identifier of this terminology database | Alphanumeric Code |
| <tbDescription> | Text describing this terminology database | Text |
| <domainOfConcept> | Subject field of this concept entry | Selected value related to concept |
| <conceptLastModified> | Date that information pertaining to this concept was last changed | Date |
| <languageCode> | Language in which the term is used | Value selected from ISO639-1 represented in the language of the term |
| <termDefinition> | Definition of the term | Text |
| <termString> | The term itself | Text |
| <wordClass> | Grammatical class of the term | Typically noun represented by n |
| <wordGender> | Grammatical gender of the term | *Masculine* represented by m, *feminine* represented by f, or *neuter* represented by n |

| | | |
|---|---|---|
| <termLastModified> | Date that information pertaining to this term was last changed | Date |
| <usedIn> | Country in which this term is used in this language | Value selected from ISO3166 represented as an English text descriptor |

Other XML elements represent containers for this information.

**Note:** in the above sample, the implication of the description of languageCode along with the text content of the termString and termDefinition elements means that the XML attribute *xml:lang* **should** be introduced into the markup to show, for example, that both the language code and the language used to represent this code is German, for example: `<languageCode xml:lang="de">Deutsch</languageCode>`. The introduction of this attribute **should** occur at the topmost point at which it is required to override the value of xml:lang propagated from elements higher in the structure.

#### D.3 Conformance to TMF

##### D.3.1 Meta-model specification

By comparison of the XML outline of this example with the structural nodes of the meta-model, the degree of conformance to the meta-model can be evaluated. The table below shows this comparison.

| Meta-Model Identifier | Vocabulary |
|---|---|
| TDC | <termBank> |
| GIS | |
| TE | <conceptEntry> |
| LS | <termGroup> |
| TS | |
| CI | |

For this example, there is no equivalent to the Term Section. The Term Section can, however, be introduced without loss of information. The example contains no Complementary Information, while the Global Information can be created out of the `<tbid>` and `<tbDescription>` elements. The result of these alterations is shown below. Bold XML elements denote the structural nodes, with bold italics denoting newly introduced sections. The *xml:lang* attribute has also been added, in italics, where needed.

```
<termBank xml:lang="en">
     <globalInformation>
          <tbid>00aa</tbid>
          <tbDescription>Automotive Engineering</tbDescription>
     </globalInformation>
     <conceptEntry>
          <domainOfConcept>ABS</domainOfConcept>
          <conceptLastModified>21-08-2001</conceptLastModified>
          <termGroup xml:lang="de">
               <languageCode>Deutsch</languageCode>
               <termDefinition> Bauteile, die die elektronischen Steuer- und
               Regelvorgänge für die Blockierregelung und die
               Antriebsschlupfregelung übernehmen.</termDefinition>
               <termSection>
                    <termString>ABS/ASR-Steuerung</termString>
                    <usageDescriptors xml:lang="en">
                         <usedIn>Germany</usedIn>
                         <usedIn>Switzerland</usedIn>
                    </usageDescriptors>
                    <wordClass>n</wordClass>
                    <wordGender>f</wordGender>
                    <termLastModified>21-08-2001</termLastModified>
               </termSection>
          </termGroup>
          <termGroup>
               <languageCode>English</languageCode>
               <termSection>
                    <termString>ABS/ASR control</termString>
                    <usageDescriptors>
```

```
                    <usedIn>Britain</usedIn>
                </usageDescriptors>
                <wordClass>n</wordClass>
                <termLastModified>20-08-2001</termLastModified>
            </termSection>
        </termGroup>
    </conceptEntry>
</termBank>
```

#### D.3.2 Data Category Specification (DCS)

Based on the description of content of elements given above, the following table shows example mappings of information units to data categories in ISO 12620 and to required data categories outside ISO 12620 such as ISO 639.

| DatCat: "**created by**" | |
|---|---|
| **DC identifier:** ISO12620A-10020201<br>**Level(s):** LS TCS TE TS | **Style:** ValuedElement<br>**Vocabulary:** termNote<br>**Example:** `<termNote value="DatCat Value"/>` |

| DatCat: "**date**" | |
|---|---|
| **DC identifier:** ISO12620A-100201<br>**Level(s):** | **Style:** Element<br>**Vocabulary:** date<br>**Example:** `<date>DatCat Value</date>` |

| DatCat: "**definition**" | |
|---|---|
| **DC identifier:** ISO12620A-0501<br>**Level(s):** LS TE TS | **Style:** Element<br>**Vocabulary:** termDefinition<br>**Example:** `<termDefinition>DatCat Value</termDefinition>` |

| DatCat: "**file identifier**" | |
|---|---|
| **DC identifier:** ISO12620A-1017<br>**Level(s):** GIS | **Style:** Element<br>**Vocabulary:** tbid<br>**Example:** `<tbid>DatCat Value</tbid>` |

| DatCat: "**geographical usage**" | |
|---|---|
| **DC identifier:** ISO12620A-020302<br>**Level(s):** TS | **Style:** Element<br>**Vocabulary:** usedIn<br>**Example:** `<usedIn>DatCat Value</usedIn>` |

| DatCat: "**grammatical gender**" | |
|---|---|
| **DC identifier:** ISO12620A-020202<br>**Level(s):** TCS TS | **Style:** Element<br>**Vocabulary:** wordGender<br>**Example:** `<wordGender>DatCat Value</wordGender>` |

| DatCat: "**masculine**" | |
|---|---|
| **DC identifier:** ISO12620A-02020201<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** m<br>**Example:** `m` |

| DatCat: "**feminine**" | |
|---|---|
| **DC identifier:** ISO12620A-02020202 | **Style:** Simple |

| | |
|---|---|
| **Level(s):** | **Vocabulary:** f<br>**Example:** `f` |

DatCat: **"neuter"**

| | |
|---|---|
| **DC identifier:** ISO12620A-02020203<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** n<br>**Example:** `n` |

DatCat: **"language identifier"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100701<br>**Level(s):** LS | **Style:** Element<br>**Vocabulary:** languageCode<br>**Example:** `<languageCode>DatCat Value</languageCode>` |

DatCat: **"en.639-1"**

| | |
|---|---|
| **DC identifier:** lang-1<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** English<br>**Example:** `English` |

DatCat: **"de.639-1"**

| | |
|---|---|
| **DC identifier:** lang-2<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** Deutsch<br>**Example:** `Deutsch` |

DatCat: **"part of speech"**

| | |
|---|---|
| **DC identifier:** ISO12620A-020201<br>**Level(s):** TCS TS | **Style:** Element<br>**Vocabulary:** wordClass<br>**Example:** `<wordClass>DatCat Value</wordClass>` |

DatCat: **"project subset"**

| | |
|---|---|
| **DC identifier:** ISO12620A-100303<br>**Level(s):** GIS | **Style:** Element<br>**Vocabulary:** tbDescription<br>**Example:** `<tbDescription>DatCat Value</tbDescription>` |

DatCat: **"subject field"**

| | |
|---|---|
| **DC identifier:** ISO12620A-04<br>**Level(s):** TE | **Style:** Element<br>**Vocabulary:** domainOfConcept<br>**Example:** `<domainOfConcept>DatCat Value</domainOfConcept>` |

DatCat: **"term"**

| | |
|---|---|
| **DC identifier:** ISO12620A-01<br>**Level(s):** TCS TS | **Style:** Element<br>**Vocabulary:** termString<br>**Example:** `<termString>DatCat Value</termString>` |

DatCat: **"terminology management transactions"**

| | |
|---|---|
| **DC identifier:** ISO12620A-1001<br>**Level(s):** LS TCS TE TS | **Style:** Element<br>**Vocabulary:** termManProc<br>**Example:** `<termManProc>DatCat Value</termManProc>` |

| | |
|---|---|
| | **Refinement:** <termManProcGrp><termManProc>*DatCat Value*</termManProc>...</termManProcGrp> |

| DatCat: "**modification**" | |
|---|---|
| **DC identifier:** ISO12620A-100103<br>**Level(s):** | **Style:** Simple<br>**Vocabulary:** modification<br>**Example:** modification |

| DatCat: "**approval**" | |
|---|---|
| **DC identifier:** ISO12620A-100105<br>**Level(s):** | |

| DatCat: "**withdrawal**" | |
|---|---|
| **DC identifier:** ISO12620A-100106<br>**Level(s):** | |

| DatCat: "**proposal**" | |
|---|---|
| **DC identifier:** ISO12620A-100110<br>**Level(s):** | |

Many of the information units in the example are bijective with ISO 12620 data categories. Ideally, this would be true for all information units.

There are exceptions to this rule in the example presented which need to be addressed: Firstly, the XML element <usageDescriptors> does not itself have content. For a TML, this grouping is unnecessary and hence can be dropped. Secondly, the XML elements with the suffix *LastModified* do not have direct equivalents in ISO 12620. To complete the mapping, appropriate encoding is required. *LastModified* contains a date that refers to the last time a modification was made. There are, in fact, two information units encoded here: a main unit that denotes a terminology management process, in this case a modification, has occurred, and a date on which it occurred. These two information units do map to ISO 12620. As the date is a refinement information unit to the terminological management process, this information should be grouped accordingly, for example:

```
<termManProcGrp>
     <termManProc>modification</termManProc >
     <modifiedDate>20-08-2001</modifiedDate>
<termManProcGrp>
```

The application of *brack* is appropriate to the GMT representation of this data.

#### D.3.3 Content mappings

For interoperability, where specific lists of data are expected to form the content of certain XML elements, mapping such shared identifiers can simplify these processes. As an example, consider a translation of the content of the languageCode element to a code based on ISO639-1. For example, *English* could become *en-639.1*. Similarly, country codes can be mapped to ISO3166.

#### D.3.4 TMF-conformant XML-representation (GMT)

The result of this analysis and substitution of identifiers for those in this standard and in ISO 12620 produces the following GMT formatted data which can be considered as a TMF-conformant TML.

This TML could be transformed automatically using, for example, XSLT to the formats specified in normative annexes B and C of this standard, and back, without loss of information.

```
<struct type="TDC" xml:lang="en">
     <struct type="GIS">
          <feat type="subsetIdentifier-12620A.10.3">00aa</feat>
          <feat type="projectSubset-12620A.10.3.3">Automotive Engineering</feat>
```

```
        </struct>
        <struct type="TE">
             <feat type="subjectField-12620A.4">ABS</feat>
             <brack>
                  <feat type="terminologyManagementTransactions-
                  12620A10.1">modification-12620A.10.1.3</feat>
                  <feat type="modificationDate-12620A.10.2.1.3">21-08-2001</feat>
             </brack>
             <struct type="LS" xml:lang="de">
                  <feat type="languageIdentifier-12620A10.7.1"> de-639.1 </feat>
                  <feat type="definition-12620A.5.1">Bauteile, die die elektronischen
                  Steuer- und Regelvorgänge für die Blockierregelung und die
                  Antriebsschlupfregelung übernehmen.</feat>
                  <struct type="TS">
                       <feat type="term-12620A.1">ABS/ASR-Steuerung</feat>
                       <feat type="geographicalUsage-12620A.2.3.2" xml:lang="en"> DE-
                       3166.1 </feat>
                       <feat type="geographicalUsage-12620A.2.3.2" xml:lang="en"> CH-
                       3166.1 </feat>
                       <feat type="partOfSpeech-12620A.2.2.1">n</feat>
                       <feat type="grammaticalGender-12620A.2.2.2">feminine-
                       12620A.2.2.2.2</feat>
                       <brack>
                            <feat type="terminologyManagementTransactions-
                            12620A10.1">modification-12620A.10.1.3</feat>
                            <feat type="modificationDate-12620A.10.2.1.3">21-08-
                            2001</feat>
                       </brack>
                  </struct>
             </struct>
             <struct type="LS">
                  <feat type="languageIdentifier-12620A.10.7.1"> en-639.1 </feat>
                  <struct type="TS">
                       <feat type="term-12620A.1">ABS/ASR control</feat>
                       <feat type="geographicalUsage-12620A.2.3.2" xml:lang="en"> GB-
                       3166.1 </feat>
                       <feat type="partOfSpeech-12620A.2.2.1">n</feat>
                       <brack>
                            <feat type="terminologyManagementTransactions-
                            12620A10.1">modification-12620A.10.1.3</feat>
                            <feat type="modificationDate-12620A.10.2.1.3">21-08-
                            2001</feat>
                       </brack>
                  </struct>
             </struct>
        </struct>
   </struct>
```

Data categories from ISO 12620 are denoted in bold. Bold italics denote references to other data categories such as those in ISO 639 and ISO 3166. By making reference to such stable systems, greater degrees of interoperability are assured.