INTERNATIONAL STANDARD

ISO 12616-1

First edition
2021-09

Terminology work in support of multilingual communication —

Part 1:
Fundamentals of translation-oriented terminography

Travail terminologique appuyant la communication multilingue —

Partie 1: Principes fondamentaux de la terminographie axée sur la traduction

COPYRIGHT PROTECTED DOCUMENT

Contents

Foreword

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

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 37, *Language and terminology*, Subcommittee SC 2, *Terminology workflow and language coding*.

This document cancels and replaces the ISO 12616:2002, which has been technically revised.

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

— updates to focus on the broader environment in which terminology workers operate;

— deepening of the aspect of terminological data management and addition of processes, tools and skills necessary for terminology tasks;

— updates to align with the technical state-of-art and the evolution of the profession.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Terminology work is conducted by people with different backgrounds and for different purposes. This document focuses on the fundamentals necessary to perform basic terminology work in translation contexts. While some details are occasionally given, the document provides the minimum information necessary to set up and work in the simplest form of a terminological data collection (TDC). The more complex tasks and processes performed by terminologists with more sophisticated technologies and in larger production environments will be covered in a future ISO 12616-2.

For clear communication, the title "terminology worker" has been chosen to represent anyone doing terminology work as an ancillary function of their professional activities. A terminology worker might be a translator, project manager or technical communicator, and might work as a single-person enterprise, for a language service provider, or in-house at a company or other organization. Terminologists and terminology workers share the same basic skill set covered in this document; however, terminologists have broader knowledge and competences, which will be discussed further in a future ISO 12616-2.

One of the most common scenarios for a terminology worker in translation contexts is the following: a client produces documentation in a particular subject field in a source language and asks a translator to translate a variety of interrelated documents. Since no terminology was provided, the translator recognizes that it would be beneficial to document the terminology found during translation work to maintain consistency across documents in the target language. This document provides terminographical best practices and data modelling principles to this end.

In this document's examples, designations and other text elements are indicated by double quotation marks, whereas objects, concepts, properties, characteristics, and types of characteristics are indicated by single quotation marks. When referring to fields in a terminology database, the first letter of the field name is capitalized and the field name is followed by the word "field" (e.g. Term field, Transfer comment field). Data categories are indicated by slashes preceding and following the name of the data category (e.g. /term/, /transfer comment/). This markup is intended to facilitate the distinction between references to the three terminological levels and other text throughout this document.

Terminology work in support of multilingual communication —

Part 1:
Fundamentals of translation-oriented terminography

1 Scope

This document specifies requirements and recommendations related to fundamentals of translation-oriented terminography for producing sound bilingual or multilingual terminology collections. It deals with the main tasks, skills, processes and technologies for translation-oriented terminography practiced by terminology workers who do terminology work in low-complexity settings as part of non-terminological activities. It does not cover terminology management involving sophisticated workflows, a multitude of roles, or advanced terminological skills and competences.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 704, *Terminology work — Principles and methods*

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

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

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

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

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

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

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

3.1
terminology
set of *designations* (3.7) and *concepts* (3.4) belonging to one *subject field* (3.2) or subject

[SOURCE: ISO 1087:2019, 3.1.11, modified — "domain" replaced with "subject field", which is the preferred term in this document.]

3.2
subject field
domain
field of special knowledge

Note 1 to entry: The borderlines and the granularity of a subject field are determined from a purpose-related point of view. If a subject field is subdivided, the result is again a subject field.

[SOURCE: ISO 1087:2019, 3.1.4, modified — The preferred term in this document is "subject field", instead of "domain".]

3.3
object
anything perceivable or conceivable

Note 1 to entry: Objects can be material (e.g. 'engine', 'sheet of paper', 'diamond'), immaterial (e.g. 'conversion ratio', 'project plan') or imagined (e.g. 'unicorn', 'scientific hypothesis').

[SOURCE: ISO 1087:2019, 3.1.1]

3.4
concept
unit of knowledge created by a unique combination of characteristics

Note 1 to entry: Concepts are not necessarily bound to particular natural languages. They are, however, influenced by the social or cultural background which often leads to different categorizations.

[SOURCE: ISO 1087:2019, 3.2.1 modified — Removed Note 2 to entry.]

3.5
individual concept
concept (3.4) that corresponds to a unique *object* (3.3)

EXAMPLE   'Saturn', 'Eiffel Tower', 'Moon', 'serial number FRHR603928', '2016 Nobel Prize in Physics'.

Note 1 to entry: Individual concepts are represented by *proper names* (3.10).

[SOURCE: ISO 1087:2019, 3.2.8]

3.6
general concept
concept (3.4) that corresponds to a potentially unlimited number of *objects* (3.3) which form a group by reason of shared properties

EXAMPLE   'planet', 'tower', 'moon', 'Nobel Prize in Physics'.

Note 1 to entry: For a general concept, it is essential that a number of corresponding objects greater than 1 can be perceived or conceived of. For example, 'spaceship' has been a general concept before such a material object existed, at the time when there existed only 1 such object, and later, when there existed several such objects.

[SOURCE: ISO 1087:2019, 3.2.9]

3.7
designation
representation of a *concept* (3.4) by a sign which denotes it in a *subject field* (3.2) or subject

Note 1 to entry: A designation can be linguistic or non-linguistic. It can consist of various types of characters, but also punctuation marks such as hyphens and parentheses, governed by domain-, subject-, or language-specific conventions.

Note 2 to entry: A designation can be a *term* (3.8) including *appellations* (3.9), a *proper name* (3.10), or a *symbol* (3.11).

[SOURCE: ISO 1087:2019, 3.4.1; modified — Removed "designator" as a synonym. In the definition, "domain" replaced with "subject field".]

3.8
term
designation (3.7) that represents a *general concept* (3.6) by linguistic means

EXAMPLE   "laser printer", "planet", "pacemaker", "chemical compound", "¾ time", "Influenza A virus", "oil painting".

Note 1 to entry: Terms may be partly or wholly verbal.

[SOURCE: ISO 1087:2019, 3.4.2]

3.9
appellation
term (3.8) that is applied to a group of *objects* (3.3) whose relevant properties are identical

EXAMPLE   "Nokia 7 Plus®" (mobile phone), "Adobe® Acrobat® X Pro" (software), "Road King®" (motorcycle)[1].

[SOURCE: ISO 1087:2019, 3.4.3]

3.10
proper name
designation (3.7) that represents an *individual concept* (3.5)

EXAMPLE   "International Organization for Standardization", "IBM®"[2], "British Isles", "United Nations".

[SOURCE: ISO 1087:2019, 3.4.4]

3.11
symbol
designation (3.7) that represents a *concept* (3.4) by non-linguistic means

Note 1 to entry: There are several types of symbols such as graphical symbols [ISO 3864 (all parts)] and letter symbols [ISO 80000 (all parts)].

[SOURCE: ISO 1087:2019, 3.4.5]

3.12
terminology work
terminology management
work concerned with the collection, description, processing and presentation of *concepts* (3.4) and their *designations* (3.7)

Note 1 to entry: Terminology work often aims at creating and maintaining *terminological data collections* (3.21).

Note 2 to entry: Terminology work often aims at terminology planning and can involve all of concept harmonization, term harmonization, and term formation.

Note 3 to entry: Terminology work can be carried out in a systematic or an ad hoc fashion.

[SOURCE: ISO 1087:2019, 3.5.1, modified — In the definition, removed "systematic". In Note 1 to entry, "terminology resources" replaced with "terminological data collections". Added Note 3 to entry.]

3.13
prescriptive terminology work
terminology work (3.12) that aims at deciding on preferred usage of *designations* (3.7)

3.14
descriptive terminology work
terminology work (3.12) that aims at documenting *designations* (3.7) as they are used in contexts without favouring preferred usage

---

1)   Nokia 7 Plus® is a trademark of Nokia Corporation, Adobe® Acrobat® X Pro is a trademark of Adobe Systems, Road King® is a trademark of Harley-Davidson. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the products named.

2)   IBM® is a trademark of International Business Machines Corporation. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named.

3.15
terminography
terminology work ([3.12](#)) aimed at creating and maintaining *terminological data collections* ([3.21](#))

[SOURCE: ISO 1087:2019, 3.5.2, modified — "terminology resources" replaced with "terminological data collections".]

3.16
term extraction
terminology work ([3.12](#)) that involves the identification and excerption of *terminological data* ([3.19](#)) by searching through a *text corpus* ([3.17](#))

Note 1 to entry: *Terminological data* ([3.19](#)) of primary interest are typically *designations* ([3.7](#)), definitions and contexts.

Note 2 to entry: Term extraction is often supported by dedicated software tools.

[SOURCE: ISO 1087:2019, 3.5.6]

3.17
text corpus
corpus
collection of natural language data

Note 1 to entry: Text corpora can be used for various activities such as text analysis or *terminology work* ([3.12](#)).

[SOURCE: ISO 1087:2019, 3.6.4]

3.18
candidate term
string of characters that has been collected by means of *term extraction* ([3.16](#)) but has not yet been selected as a text element to be documented in the *terminological data collection* ([3.21](#))

3.19
terminological data
data related to *concepts* ([3.4](#)) and their *designations* ([3.7](#))

Note 1 to entry: Common terminological data include *designations* ([3.7](#)), definitions, contexts, notes to entry, grammatical labels, subject labels, language identifiers, country identifiers, and source identifiers.

[SOURCE: ISO 1087:2019, 3.6.1]

3.20
terminological entry
concept entry
CE
collection of *terminological data* ([3.19](#)) related to only one *concept* ([3.4](#))

[SOURCE: ISO 1087:2019, 3.6.2, modified — Added "concept entry" and "CE" as preferred and admitted terms, respectively.]

3.21
terminological data collection
TDC
terminology resource
resource consisting of *concept entries* ([3.20](#)) with associated metadata and documentary information

[SOURCE: ISO 26162-1:2019, 3.2.4, modified — Added "terminology resource" from ISO 1087:2019, 3.7.1.]

3.22
data category
class of data items that are closely related from a formal or semantic point of view

EXAMPLE   /part of speech/, /subject field/, /definition/.

Note 1 to entry: A data category can be viewed as a generalization of the notion of a field in a database.

Note 2 to entry: In running text, such as in this document, data category names are enclosed in forward slashes (e.g. /part of speech/).

[SOURCE: ISO 26162-1:2019, 3.2.11]

3.23
terminology management system
TMS
software tool specifically designed with a metadata structure for collecting, maintaining, and accessing *terminological data* (3.19)

[SOURCE: ISO 1087:2019, 3.6.13]

3.24
concept orientation
principle whereby a *concept entry* (3.20) describes a single *concept* (3.4)

Note 1 to entry: When two or more different *concepts* (3.4) are represented by the same *designation* (3.7) (in the same language), this designation is considered a homograph. Such *concepts* (3.4) are documented in separate *concept entries* (3.20).

[SOURCE: ISO 26162-1:2019, 3.2.13]

3.25
term autonomy
principle whereby all *terms* (3.8) in a *concept entry* (3.20) are considered independent sub-units and can be described using the same set of *data categories* (3.22)

Note 1 to entry: By analogy, this principle applies to *designations* (3.7) as well as other text elements.

[SOURCE: ISO 26162-1:2019, 3.2.14]

3.26
data granularity
degree of precision of data

Note 1 to entry: For example, the set of individual *data categories* (3.22) /part of speech/, /grammatical gender/, and /grammatical number/ provides for greater data granularity than does the single data category /grammar/.

[SOURCE: ISO 26162-1:2019, 3.2.15]

3.27
repeatability
principle whereby a *data category* (3.22) can be repeated within a database definition and whereby it can also be combined with other data categories

[SOURCE: ISO 26162-1:2019, 3.2.12]

3.28
data elementarity
principle whereby a data field contains only one data element

EXAMPLE   For example, including both a full form and an abbreviation of a term in the same data field would be a violation of data elementarity.

[SOURCE: ISO 26162-1:2019, 3.2.16]

3.29
terminology worker
person whose role is to perform *terminology work* ([3.12](#)) as an ancillary function of other professional activities

3.30
terminologist
expert who performs *terminology work* ([3.12](#)) as a main function of a professional activity

3.31
technical communicator
expert who defines, creates and delivers information products for the safe, efficient and effective use of products

Note 1 to entry: Products may be technical systems, software, or services.

3.32
source language
language of the content to be translated

[SOURCE: ISO 18587:2017, 3.2.2]

3.33
target language
language into which source language content is translated

[SOURCE: ISO 17100:2015, 2.3.6]

3.34
transfer comment
note in a *terminological data collection* ([3.21](#)) providing information on the degree of equivalence, directionality or other special features affecting equivalence between a *designation* ([3.7](#)) in one language and another designation in a second language

4   Fundamentals of terminology management

4.1   Goals

Terminology management has various goals. A basic goal is that users of a terminological data collection (TDC) are able to retrieve information that answers their question. The TDC shall thus cover the pertinent subject fields and terminologies. Concept entries in the collection shall be correct and complete and not exhibit any data integrity issues with other entries in the collection.

An extension of the basic goal above for translation support is that data is optimised for use in computer-aided translation (CAT) systems. Such systems have a terminology component. During the translation process, the content of the translation segment in the source language is matched against the TDC. Matches are displayed in the terminology component or made available for easy integration into the translation in the target-language segment.

4.2   Work environments

Work in pursuit of this basic goal is performed in many different work environments with one person or several doing terminology work, supported by tools, following working methods along a particular workflow, etc. These terminology workers document the result of their work, and their work is guided by training and documentation. They have certain skills and deal with one, two or many languages.

[Figure 1](#) shows a list of features that characterise work environments where terminology work is performed. The colour intensity shows the varying degree of complexity of a feature. For example,

the top-most arrow indicates that terminology work can support one process in low-complexity environments and several processes in more complex environments.
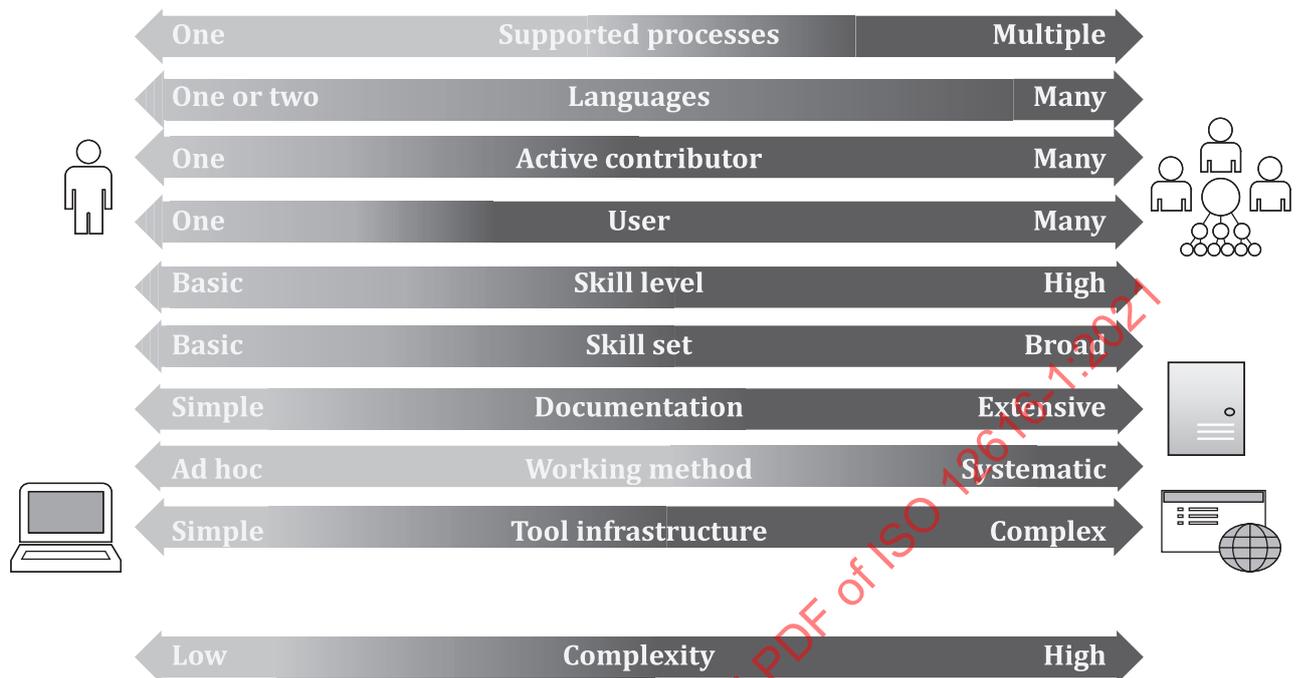
One	Supported processes	Multiple
One or two	Languages	Many
One	Active contributor	Many
One	User	Many
Basic	Skill level	High
Basic	Skill set	Broad
Simple	Documentation	Extensive
Ad hoc	Working method	Systematic
Simple	Tool infrastructure	Complex

Low	Complexity	High

Figure 1 — Level of complexity in different environments

The type or complexity of the feature varies from one environment to the next or even from one project within the environment to the next project, as the colour shading indicates.

EXAMPLE 1   A freelance translator who works on terminology for a patent translation project is the only active contributor to the TDC and the only user of the data. Although he does not do terminology work full time, he can have a very high level and broad set of terminology skills. And for certain concepts, he might need to do extensive research.

EXAMPLE 2   In contrast, a terminologist works in a complex tool infrastructure. She is one of many terminologists who serve dozens of languages. In one project, she might be doing only ad hoc terminology work (see 4.4.2) to try to solve a particular terminological problem quickly.

Much of this document applies to any environment. The main focus, however, is on the low-complexity types of environments, which can be characterised as follows.

— One process (e.g. support of the translation process) is the main focus.

— The need for documentation, e.g. in the form of a guide or training material, is low.

— The terminological data collection itself is simpler and generally contains fewer data categories.

— There is one active contributor or only a few. They can include the translator, a subject-matter expert, and the client. And there might only be a few users. That means that the workflow is simple.

— The skill level is assumed to be low and fairly narrow. This document covers anything that is necessary to set up a correct terminological entry with a minimum of terminological information.

— This document can be helpful in a monolingual environment. But terminology work for at least two languages is assumed.

— Ideally, a terminology worker uses a terminology management system (TMS), but many aspects of this document also apply to simpler repositories (e.g. spreadsheets).

4.3 Translation vs. terminology work

Terminology work is not to be confused or equated with the process of translation, which is concerned with rendering source language content into target language content. The translation process typically involves two languages. Terminology work can be monolingual, bilingual, or multilingual. It is focused on concepts and their designations in a specific subject field. While terminology work aims at understanding and describing concepts in a way that is independent of individual contexts of communication, translation is always dependent on contextual and situational meaning.

Similarly, a terminology database is different from a translation memory. Both are databases, but the former is focused on managed concepts and contains their designations plus supporting terminological information, while the latter contains translation segments in two languages. A translation memory can at times serve as a resource for terminology work, e.g. as a text corpus for bilingual term extraction.

In the context of translation-oriented terminology work, translators or terminology workers need to determine target-language equivalents for designations used in the source text. These source-language designations shall not be "translated"; on the contrary, target-language equivalents shall be established based on conceptual equivalence. This shall be done independently in each of the target languages involved in the translation project. Terminological entries resulting from this research should provide equivalents not just for the current project, but for future translation projects from and into all of the working languages covered in the terminological entries (see 8.2.2).

4.4 Working methods

4.4.1 Conceptual analysis

Multilingual and translation-oriented TDCs shall provide translators with designations in source and target languages that are considered equivalents referring to one and the same concept (see 8.2.2). When recording equivalents in a TDC, a terminology worker shall not rely solely on linguistic similarities between designations (e.g. similar morphology, similar type of motivation). Instead, the terminology worker shall perform a conceptual analysis of relevant material offering information on the concepts in question. Such an analysis shall apply the principles laid out in ISO 704.

To do that, the terminology worker needs to be aware of the basic theoretical model supporting any terminology work: human beings perceive objects of their physical world or conceive objects in their thinking. Perceiving or conceiving one or more objects, the relevant properties of these objects are abstracted into characteristics and form a concept. Concepts are thus purely mental representations and correspond to the objects in question. For human communication about objects or their concepts, various types of designations (e.g. terms and proper names) are used. To be clear about the concepts, definitions are used.

EXAMPLE   Office environments involving computers include objects referred to as 'computer mouse' or 'mouse'. Each of these objects has individual properties, such as 'has two buttons' or 'has three buttons', 'performs two-dimensional rolling on a flat surface', 'has a cable connection' or 'has a wireless connection', 'used for moving a cursor on a display', etc. The relevant properties are abstracted into characteristics and form a concept that can be defined as follows: 'computer input device having one or more buttons and capable of two-dimensional rolling motion which can drive a cursor on the display and which performs a variety of selection options or commands', (ISO 9241-400:2007, 3.6.11). Depending on the given communication purpose and situation, the term "computer mouse" or "mouse" can be used for communicating about this concept.

In terminology work, conceptual analysis involves all steps necessary to identify and clarify concepts. A terminology worker needs to cognitively reconstruct the process of concept formation. This involves identifying essential characteristics, comparing the concept with neighbouring or related concepts, and delineating the concept from these concepts. In the Example above, this would mean understanding how to differentiate between a 'computer mouse' and other pointing devices. In multilingual and translation-oriented terminology work, a terminology worker shall verify whether in all involved languages the set of essential characteristics (the intension) is the same.

It is important to note that concepts should be viewed and analyzed in their conceptual context. The result of such analysis is a concept field or concept system, which in turn can be depicted by means of

non-formalised concept diagrams (see ISO 704:—, 5.5.4.2.1, Example 1) or formalised concept models (see ISO 24156-1).

In the Example above, conceptual analysis would lead to the understanding that 'computer mouse' is a specific concept of 'pointing device.' In the definition, this relation is indicated by the reference to 'computer input device'. Further conceptual analysis might reveal whether specific concepts of 'computer mouse' exist or of which partitive concepts the concept 'computer mouse' is made of.

Comprehensive conceptual analysis is time-consuming, but mandatory in systematic terminology work (see 4.4.2). In translation-oriented terminology work, conceptual analysis may therefore be limited to what is necessary to ensure understanding of concepts and determining equivalence in a given translation project.

4.4.2   Systematic vs. ad hoc terminology work

While terminologists engaged in classification of subject fields mainly use a systematic approach to organize the concepts in the subject field, terminology workers in translation environments use an ad hoc approach in most cases. Table 1 juxtaposes some of the critical differences.

Table 1 — Systematic vs. ad hoc terminology work

	Systematic	Ad hoc
Focus	Classification of the concepts in a subject field	Solving terminological problems that occur in specific situations, e.g. a translation project
Linguistic level	Designations only	Designations and other text elements (see Figures 3 and 4)
Main approach	Onomasiological	Semasiological
Starting point	One subject field	One text in a source language
Result	Concept entries that are more or less connected in a larger system	Individual entries

A terminology worker performing terminology work in preparation of a translation project will often find textual segments that require further research. Such a segment can turn out to be

— a designation representing a clearly delineated concept (see 7.2), or

— some other elements (see 7.3).

In most cases, the text element can very easily be explained through an ad hoc approach, e.g. with a search in an online bilingual dictionary or text corpus and the subsequent simple documentation in the TDC. The terminology worker shall be aware that some translation problems can only be solved through conceptual analysis (see 4.4.1) in the various languages and the systematic approach where multiple concepts are compared and organized according to how they relate to each other.

4.4.3   Prescriptive vs. descriptive terminology work

Terminology workers shall be aware of the following three approaches to terminology work:

— descriptive;

— prescriptive;

— normative.

The goal of descriptive terminology work is to document designations as they are used without favouring preferred usage. The goal of prescriptive work is to establish rules of term usage, e.g. by setting the respective usage status for a term in the TDC. Only terminology workers in standardizing

bodies do a form of prescriptive terminology work called normative. Terminologies established in such settings become normative.

EXAMPLE 1   A small group of translators working on the same project need to use terminology consistently. One person might collect designations found in existing text corpora (descriptive). The group might then decide which designations are preferred or admitted and which should be avoided in the texts of the client or sponsoring organization. Therefore, the work would be prescriptive even if the preferred forms are not labelled as such in the TDC.

EXAMPLE 2   A client provides a list of preferred designations to a translator.

EXAMPLE 3   A committee of subject-matter experts working on standards in their subject field define the terminology and use it in such standards. Standards are normative documents and can be useful reference material for translators.

When consulting a TDC in their own research, terminology workers shall be aware of the approach taken in the creation of the TDC in question.

4.4.4   Text corpora

Terminology workers benefit from the availability of text corpora for their work. A collection of texts in machine-readable form can take as simple a form as documents in a folder on the terminology worker's computer. Text corpora can be (see Reference [16]):

— monolingual, bilingual or multilingual;

— parallel or comparable;

— specific to one or more subject fields;

— pre-existing or created for the project at hand;

— non-web or web;

— unprocessed or pre-processed (e.g. annotations, deletions);

— static or dynamic (i.e. texts are added or removed over time);

— freely available or paid.

Text corpora are useful to answer questions on all levels of the semantic triangle (see ISO 704):

— the object (e.g. how does this piece of equipment work?);

— the concept (e.g. what exactly is this?); and

— the designation (e.g. how do experts refer to the concept in the target language?).

Evaluating designations in their natural linguistic environment provides answers to questions, such as the following:

— Is the designation used?

— Is it used widely?

— Who uses it?

— What words or designations co-occur with the designation in question?

— What synonyms to the designation exist?

— What spelling variants are used most frequently?

5 Process steps and activities

5.1 Setting goals

Translation-oriented terminology work is frequently part of a clearly defined project with a beginning and an end. Terminology work is a process following a sequence of steps. Not all steps are necessary in every project, but it is most efficient to follow a sequence.

At the beginning of a terminology project in support of translation work, project goals shall be defined, even if only in an informal way. A project goal could be as simple as "document the most important concepts for the translation project in a prescriptive way for two languages in a spreadsheet." It could also be "spend five hours identifying designations that occur in provided documents, set up entries following the prescriptive approach in a specific terminology management system, and prepare a concept system to enable discussions with subject-matter experts."

The following aspects shall be considered.

— Languages: The languages involved in the project.

— Scope: One-off project vs. establishment of a long-term terminological data collection.

— Information needs: The data categories necessary for the project (see Clause 8).

— Timing: Terminology work completed before translation starts (proactive terminology work) vs. terminology work after translation has already started or is indeed finished (reactive terminology work).

— Term usage: In prescriptive terminology work, the terminology worker shall identify which synonym is preferred over another.

— Deployment: Means by which terminological data will be distributed in the subsequent process and its likely use (e.g. by a translator using computer-aided translation tools) (see 5.7).

— Data security: Sensitive terminology can be an issue (e.g. product names), and confidentiality shall be considered in documenting and sharing such terminology.

5.2 Basic workflow

Terminology work as a process is driven by the defined goals (see 5.1). It can be broken down into steps, as Figure 2 depicts. A terminology worker should use a terminology management system (for alternative tools, see Clause 9). This workflow implies the use of such a system. Therefore, initially, a terminology database shall be set up. That is a one-time step. Once that is in place, terminology work follows through these steps: collection of terminological data, research, documentation and use. The maintenance step is necessary when a terminological entry has to be changed.
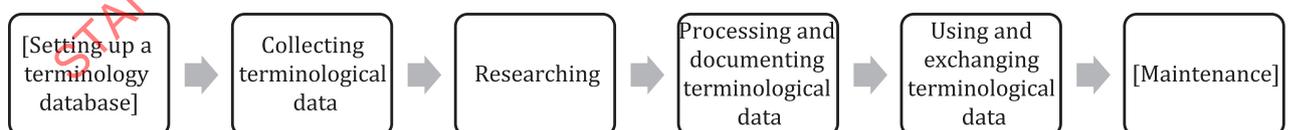
Figure 2 — Process steps

5.3 Setting up a terminology database

The goals of the terminology project determine the storage type, the structure and the content of the terminological data collection. Before a terminology database or another type of TDC is set up, terminology workers shall analyse the information needed. For example, they should look at what type of terminological data is available, what has been most helpful in the past, and what would be good to add going forward. Annex A contains a set of data categories that have proven beneficial in translation-

oriented environments. Ideally, the metadata for the TDC are selected and the terminology database is defined only once, following the principles given in 8.2, and used for all terminology management needs going forward.

A terminology management system should be used to create a TDC. The TMS should be compatible with other tools (see Clause 9) and support automation for as many aspects of the process as possible (e.g. import/export of existing data rather than manual entry; automated distribution of terminology).

Instead of a terminology management system, a simple application, such as a spreadsheet program, can be chosen if the content (number of languages and concept entries) is not expected to be very large, and if the structure (data categories, relations between entries) of the TDC is not expected to be very complex (see also Clause 9).

5.4 Collecting terminological data and initial assessment

In translation-oriented terminology work, terminology and terminological data must be collected in the source language, or the target language, or both. In most scenarios, the starting material is a document in the source language. A terminology worker has the choice of collecting data manually by going through the document and marking text elements to research. This approach is generally best, if the size of the text is small.

This step can also be supported by a term extraction program (see Table 3), which identifies candidate terms and presents them in a list. The terminology worker shall then decide whether these candidate terms are meaningful text elements or whether they are "noise", that is, strings that are not terminologically relevant. For example, text elements, such as "this is a" or "and subsequently" might occur with a high frequency, but are not related to the subject matter of the text. Such noise shall not be processed further by the terminology worker. In manual term extraction, noise rarely plays a role, as the terminology worker has context and selects only meaningful text elements to begin with.

The terminology worker shall strive for a clean list of text elements that need further examination.

5.5 Researching

Translation-oriented terminology work shall ensure both an adequate understanding of the concepts in the context of a given translation project as well as linguistically-correct usage of designations and terminologically relevant text elements. To that end, terminology workers shall take the necessary research steps. Not every text element needs to be examined equally extensively. But term assessment and concept research generally necessitate the use of resources, such as existing TDCs, style guides, subject-matter resources, experts, etc.

Term assessment begins with the separation of designations from other text elements (see 7.1). Designations should subsequently be analyzed for their type (term, proper name, symbol and name-like designation), for their part of speech and for their agreed criteria related to usability and acceptability. Where applicable (e.g. when a translation shall be terminologically consistent), terminology workers shall identify one synonym as preferred term (out of the set of synonymous terms). For term assessment, terminology workers shall follow the guidance given in ISO 704.

Other text elements may be rated according to pragmatic aspects, such as consistency and visibility.

Concept research shall result in terminological data that establishes the concept as clearly as possible, and ideally leads to a definition (see 6.3). For conceptual analysis, terminology workers shall follow guidance given in ISO 704.

Depending on the complexity of the concept, researching and confirming the designations can be completed fairly quickly, in one step, or can take longer and involve several steps. Ideally, related terms and concepts should be researched together, both to highlight their semantic differences and for efficiency purposes.

Bilingual dictionaries are a useful resource for finding term equivalents in different languages. However, like terms found in translated documents, all terms found in bilingual dictionaries should be confirmed by determining their use in reliable monolingual documents.

5.6 Processing and documenting terminological data

The terminology worker should process the information gained during the research phase and document it in the terminological data collection. This step includes:

— integrating the newly researched information into the TDC (e.g. setting up a new entry and linking it to existing entries);

— following terminographical best practices (see Clause 6).

In addition to the points made in Clause 6, terminographical best practices include:

— filling in mandatory and optional data categories (see 8.4), as necessary;

— ensuring that all synonyms are appropriately documented in concept entries;

— indicating the usage of designations and other text elements to facilitate translation or the drafting of texts;

— indicating the workflow status of concept entries;

— avoiding the creation of duplicate concept entries (doublettes); for dealing with duplicate entries, see 5.8.

5.7 Using and exchanging terminological data

If all preceding steps were carried out or applied correctly following the principles laid out in Clause 6, the following will occur: During the translation process in the translation environment, when a documented source term appears in the source text, the terminology component of the translation tool makes the term in the target language available to the translator. The success of such matching processes depends on a variety of factors:

— the preparation of designations and other text elements in their base form (see 6.2);

— the adherence to the principle of data elementarity in the Term field (see 8.2.4);

— the matching functionality of the tool;

— the complexity of the language with regards to inflection, etc.

Ideally, the TMS is integrated in the translation environment. Alternatively, the TMS and the translation environment can be connected via tools, such as an API or a connector. If the TDC was prepared in a spreadsheet, the spreadsheet needs to be imported into the translation tool. To enable such a data exchange, the terminology worker shall set up the spreadsheet accordingly. Annex B provides some guidance. Terminology workers shall follow the guidance provided in ISO 30042 (see also 8.1 and 8.5) to facilitate export and import procedures of terminological data from and into TDCs and especially tasks related to the exchange of terminological data.

5.8 Maintenance

During the use of the data, the need for maintenance can arise, and the terminology worker shall carry out the necessary changes in the TDC.

For example, there can be changes in the concepts or objects in question (due to technological progress or innovation in the subject field); there can be changes in term usage; or there can be a spelling reform in one of the languages. In addition, a lack of data integrity within the entry or with other entries can become evident and trigger maintenance.

A required change can impact only one entry, or it can affect multiple entries. It can be done quickly or take a significant amount of time. It can involve just revising or deleting a terminological entry entirely, or it can have ramifications that affect multiple languages and thus require the assistance of multiple terminology workers.

Designations that have become deprecated shall not be deleted but shall be kept in the terminological entry and be marked as deprecated. When duplicate entries (doublettes) are identified, one of the entries shall be deleted after ensuring that all the necessary information is stored in the retained entry.

6 Terminography

6.1 Best practices for open data categories

When recording terminological data for specific data categories, terminology workers shall follow terminographical best practices and conventions. In open data categories, content can vary (see 8.3). Guidance for the most important open data categories is provided in 6.2 to 6.8 (see also References [17] and [21]).

As a general rule, in open data categories, consistency in style shall be maintained, and grammar, punctuation as well as typography rules shall be followed (e.g. remove leading or trailing spaces when inserting terminological data).

6.2 Term

The Term field shall contain one designation only. If synonyms (e.g. a full form and its short form) exist, each shall appear in its own Term field. Otherwise, term autonomy and data elementarity are violated (see 8.2.3 and 8.2.4).

All identified synonyms, including full form, short form, orthographic variants and deprecated designations, shall be documented in the same terminological entry.

Designations shall be recorded in their base form. The base form, also known as the canonical form, is the word form chosen according to grammatical conventions to represent all possible forms of the word. For example, "bind" is the base form for the following word forms: bind, bound, binds, binding. Table 2 describes what the base form is for the most important parts of speech in English.

Table 2 — Part of speech and base form

Part of speech	Base form characteristics	Examples
Noun	Lower case, singular, unless it is a mass or plural noun	*monitor*, not *Monitors*
Verb	Lower case; without "to"; in the infinitive	*chase*, not *to chase*
		stack, not *stacked*
Adjective	Lower case	*fireproof,* not *Fireproof*
Proper name; appellation	Specific letter casing; sequence of elements in spoken order	*Dalmatian*, not *dalmatian*
		The Netherlands, not *Netherlands, The*
		Airbus A380, not *A380 by Airbus*
Adverb	Lower case	*hereinafter*

6.3 Definition

A definition is a representation of a concept by an expression that describes and differentiates it from related concepts (ISO 1087:2019, 3.3.1). Definitions are the result of conceptual research and analysis (see 4.4.1). In most TDCs, definitions serve as the main source of information about the concept.

During the research phase (see 5.5), the terminology worker should select definitions from reliable resources that adequately explain the concept. If definitions are authored, this shall be done in line with the concept system of the culture and subject field. Definitions should follow ISO 704 guidelines. The designation shall not appear in the definition.

In translation-oriented terminology work, one definition in a language that is understood by all involved translators is often sufficient. When definitions for two or more cultures are necessary, it is generally not recommended to translate definitions. Instead, the terminology worker should research the concept and its designations in each language or culture. It can be necessary to translate definitions in specific settings, e.g. in international standardization, multilingual legislation, or in-house standardization of company-specific terminology, or when introducing new terms and loan terms to a target language where no definition exists.

6.4 Concept description

A concept description can provide a wide range of information about a concept. For example, it can indicate characteristics of the concept, clarify concept relations, or illustrate how a concept functions or what it is used for. Concept descriptions do not have a conventional format and most often take the form of running text. Concept descriptions are not considered representations of concepts and in terminology work may be used as supplementary information to the definition (see ISO 704:—, 6.6). In translation-oriented terminography, concept descriptions may also be used as an alternative to definitions and are therefore useful for terminography in fast-paced but low-complexity translation environments. If used as an alternative to a definition, the concept description shall ensure adequate understanding of the concept in the context of the given translation project.

6.5 Context

A context is an authentic text segment that contains the designation as described in the terminological entry. Text segments stored in a Context field shall illustrate the use of the designation for a given situation. Contexts may also be used to illustrate the concept or certain aspects of it. Such contexts can contribute to a better understanding of the concept.

The context shall originate from the most reliable source available and be composed in the language of the designation used. For future reference, to confirm reliability, and for copyright reasons, the source of the context shall be included in a Context source field.

6.6 Note

Notes may be used whenever the terminology worker sees the need for clarification, especially to avoid misinterpretation of information in the entry, e.g. to point out the risk of confusion with similar concepts. Notes may also be used for information enhancing the understanding of the concept. Notes may contain information on term usage (e.g. why a term is deprecated) or be a transfer comment where information on the degree of equivalence or directionality is explained.

Notes shall be kept to a minimum and only be used if the information cannot be documented in other fields. Notes to the entry should be complete sentences; they should start with a capital letter and finish with a full stop.

6.7 References and source identification

Terminological information can be found in many places: in general and specialized reference works, such as encyclopaedias and dictionaries, educational textbooks, manuals and other relevant literature, terminology or text corpus databases, and from subject-matter experts. When beginning a terminology project, it is generally best to establish a good understanding of the subject field by using general reference works, then move on to more specialized sources. A good reference librarian can be very helpful.

Information that is obtained from websites of unclear reliability shall be verified in reliable original-language reference works.

Note that translated information sources may also be used, but all information from such a source shall be confirmed in reliable original-language references. In certain environments, such as in some multilingual organizations, translated sources must be used due to organizational constraints.

Sources of terminological information shall be selected carefully to ensure that the information is accurate and reliable. The quality of the information can be judged, for example, by considering

— author's or expert's credentials,

— intended readers of the documents,

— its global or regional relevance,

— publication date,

— methodology used,

— grammar and writing style, and

— whether a document is normative or legally binding.

Suitability for purpose can be determined from the

— introduction,

— table of contents,

— index,

— overall perspective and language register,

— information about the resources, such as a blurb or an abstract, and

— external reviews, such as book or website reviews, or citations in technical or academic publications.

Unambiguous information about the sources should also be recorded in the TDC, including at least the author or originator, full title, date of publication or consultation, location, and contact information for experts consulted. Contact information should be protected according to the applicable laws for the protection of personal information.

There are various reasons for keeping good source records: the references can be retrieved and shared with others in a team, references can be checked or referred to again for further information (context, phrasing, word combinations), their reliability and the reliability of the terminological information can be assessed, and copyright can be documented.

Management of source information is important for efficiency, ease of use and research, and exchange of information. Source identification for less stable information, such as that found on websites or in advertising materials, should include the consultation date with the hyperlink or publication title, date and author information. Web pages are updated frequently, and the material might or might not be archived, making retrieval difficult or impossible. It is often useful to print or make an electronic copy of such information. Source codes are often used in concept entries for convenience, with detailed source information maintained in a bibliographic database. A standard system should be used (see ISO 12615 and ISO 690), especially when data is to be exchanged with others.

6.8 Abbreviations for names of languages and countries

In translation-oriented terminography, it is often necessary to record information related to languages and countries. It is recommended to use abbreviations for the names of languages and countries, in

which case the terminology worker shall use the language codes given in ISO 639 (all parts) and the country codes given in ISO 3166 (all parts).

7 Text elements for the Term field

7.1 General

At the beginning of every translation-oriented terminology project, terminology workers shall identify text elements that need to be documented in the Term field of their TDC. Terminology workers shall be acquainted with selection criteria for such elements. Criteria and their weight vary depending on the subject field, language, project, knowledge of the terminology worker, the target group of the TDC, etc. Questions, such as the following, can help in determining how relevant a text element is.

— Do I recognize the element in the source language?

— Do I know how to express the concept in the target language?

— Where (in a text, on a user interface, etc.) does the element occur?

— How often does the element occur?

— Does it represent an important concept in the subject field?

— Is it a highly visible name?

— Does it have unusual spelling?

— Who uses the element?

— Has it been vetted by experts?

While it is helpful and recommended to establish and follow a framework of selection criteria, terminology workers have considerable freedom, especially if they are the sole users of the TDC. In systematic terminology work, only designations representing concepts are recorded, with terms being the most common. Figure 3 provides a concept diagram of various types of designations that are typically documented in the Term field. The concept diagram is not comprehensive. For example, there are more types of proper names than shown. Subclause 7.2 provides more details.
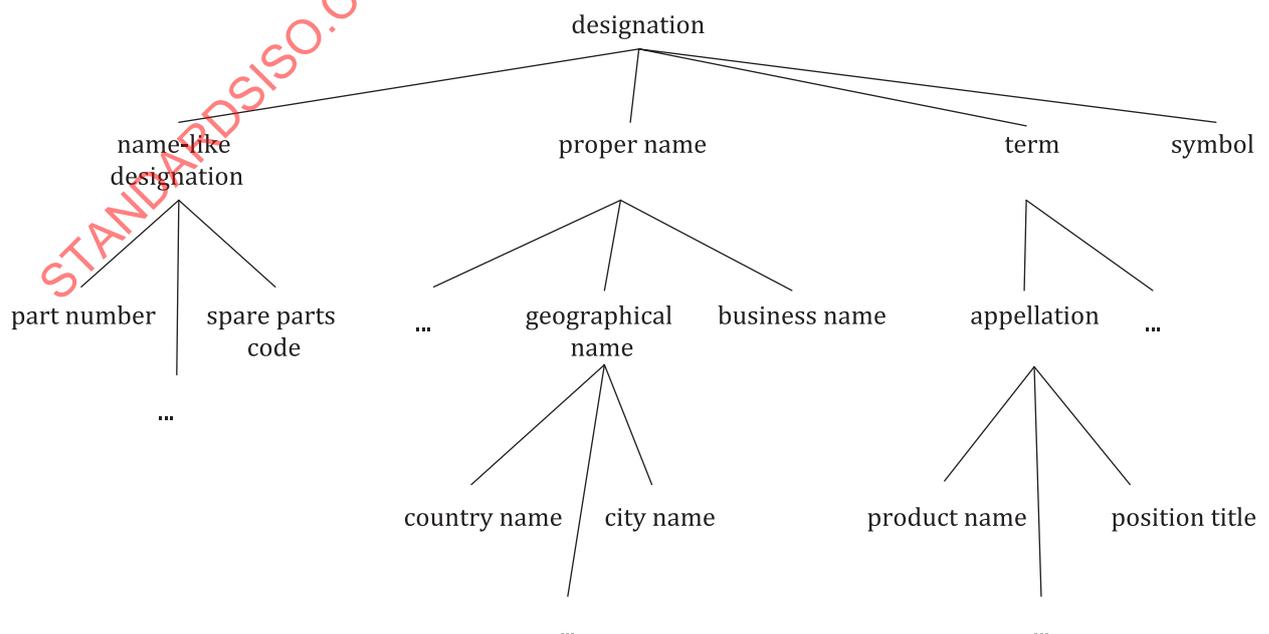
Figure 3 — Content for the Term field: designations

In today's fast-paced translation environments where terminology workers must minimize manual work, however, more pragmatic criteria may also be applied to identify other text elements that should be included in a TDC. They are shown in Figure 4 and further explained in 8.3.

Such elements do not necessarily represent concepts and do not always belong to one subject field. It is important to note that some elements can clearly represent a concept to a speaker of one language and not to a speaker of another language, since concept systems are not entirely language- or culture-independent. In these cases, a pragmatic solution should be found.
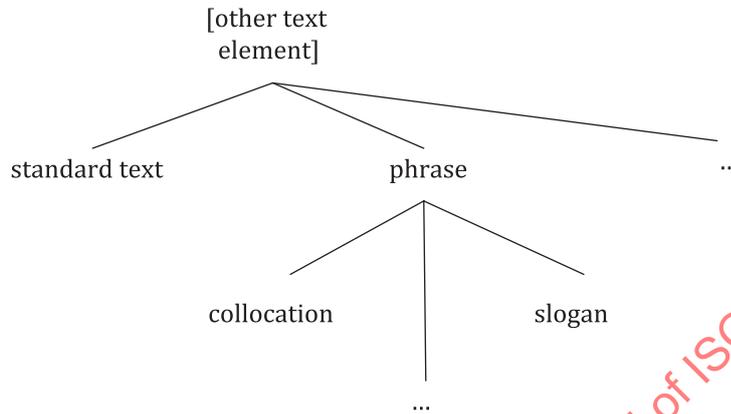

Figure 4 — Content for the Term field: other text elements

7.2   Designations

7.2.1   Terms

Terms represent general concepts that can cover a potentially unlimited number of objects. Terms consist of one or more words or lexical units. A single-word term in one language can be expressed by a multi-word term in another language.

EXAMPLE 1   The German single-word term "Atemnot" corresponds to the English multi-word term "shortness of breath".

A special category of terms are those covering objects that are largely identical. In accordance with ISO 1087:2019, 3.4.3, they are referred to as appellations.

EXAMPLE 2   Both "Ford Model T" and "Tin Lizzie" refer to a concept that corresponds to a group of largely identical automobiles. "Ford Model T" and "Tin Lizzie" are appellations and synonyms.[3]

EXAMPLE 3   "Olympic Summer Games" is the appellation of an event that occurs every four years.

7.2.2   Proper names

A proper name designates a concept that corresponds to a unique object regardless of the physical composition of the entity or the plural form of the words that make up the name; see Reference [18].

EXAMPLE 1   The geographical name "United States of America" represents a concept that corresponds to a single entity even though it is composed of 50 states. The "University Hospital Network" is a unique entity, but is made up of three hospitals and rehabilitation and educational institutes.
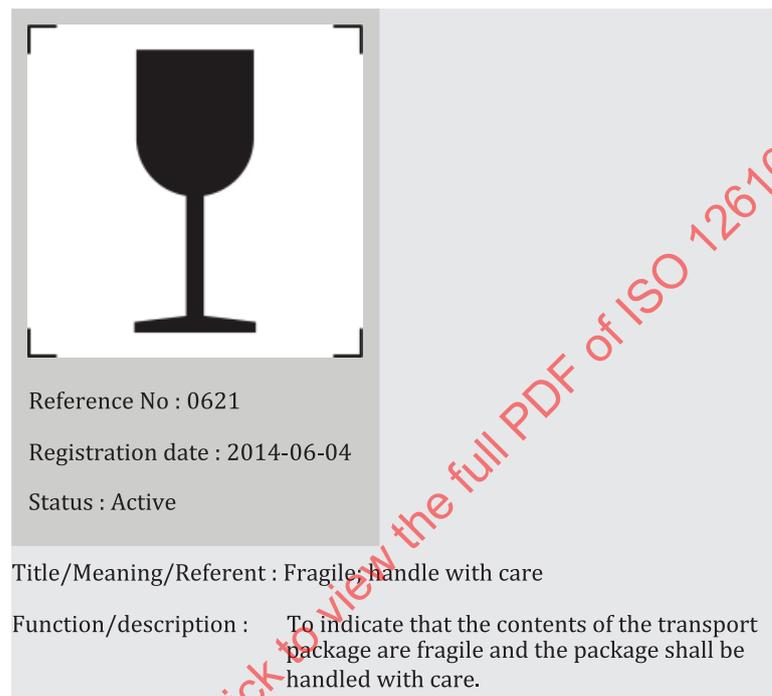
EXAMPLE 2   The "Games of the XXXII Olympiad" is the event name for a concept corresponding to a single instance of the "Olympic Summer Games".

---

3)   Ford Model T is a trademark of Ford Motor Company. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the products named.

7.2.3   Symbols

Symbols are an important aid to international communication because their visual representation of concepts often functions independently of any given language. They communicate information directly, i.e. without the use of words, under specific circumstances (e.g. traffic signs, airport signs, packaging symbols). For information about iconic, non-iconic and other types of symbols see ISO 704:—, 7.5).

Only symbols that represent a concept and can be replaced by a term or name should be considered designations. Figure 5 shows an excerpt of ISO 7000. It illustrates that the depicted symbol is a synonym to "fragile" and "handle with care".

Reference No : 0621

Registration date : 2014-06-04

Status : Active

Title/Meaning/Referent : Fragile; handle with care

Function/description :   To indicate that the contents of the transport package are fragile and the package shall be handled with care.

NOTE   The figure is included as an example of a reliable source that documents that the depicted symbol is used as a synonym to "fragile" and "handle with care." It is not an example of how synonyms shall be recorded in a TDC (see 8.2.3 and 8.2.4).

Source: ISO 7000:2019, registered symbol ISO 7000-0621.

Figure 5 — Example of a symbol as a synonym to text elements

Whenever the technology allows for their integration in terminological data collections, symbols should be added as synonyms for designations or other terminologically relevant text elements (see 7.1), for example, by inserting or attaching a graphics file. This requires, however, that the data type of the field supports this feature. If it is not possible to record symbols in a Term field, other fields may be used.

7.2.4   Name-like designations

Name-like designations, such as article numbers, parts codes, vehicle identification numbers or stock numbers, play an important role in communication about goods and products. The terminology worker may include such designations in the TDC. Such codes often convey information relevant to understanding the concept.

EXAMPLE   The Stock Number "5821-12-120-84" has, in the context of NATO, been used to identify "stock group 58 – communication equipment", "Stock class 21 – radio and television communication equipment, airborne", "Nation code 12 – Germany" and "Serial number 12-84-8477;" see Reference [19].

7.3 Other text elements

7.3.1 Phrases

In translation-oriented terminology work, besides designations, phrases of all kinds can become relevant units of information. Phrases can contain one designation, more than one designation or no designations at all. They have a fixed lexical composition and syntactic stability. Terminology workers may add them to terminology databases. Such entries contribute to enhancing the idiomaticity and linguistic consistency of a translation. The terminology worker may document such phrases and their equivalents in other languages of the translation project in the TDC.

EXAMPLE 1   "this side up", phrase on a package.

EXAMPLE 2   "chair a meeting", English collocation.

EXAMPLE 3   "Citius, fortius, altius", motto of the Olympic Games.

7.3.2 Standard texts

Standard texts, also referred to as boilerplate texts, can comprise whole sentences that have set wording and are used to serve a particular function. They often occur multiple times in documentation with little to no change. The terminology worker may document standard texts and their equivalents in other languages of the translation project in the TDC.

Translators generally rely on their translation memory systems for reuse. But a standard text chunk can be treated as a text element and managed more effectively as an entry in a terminological data collection. The terminology worker might, for example, find it useful to record information about the authors, approvers, areas of application, etc.

EXAMPLE   A standard text a company uses on its website to alert users to its cookie policy as well as its consent procedure: "We use cookies to improve your website experience. To learn about our use of cookies and how you can manage your cookie settings, please see our Cookie Policy. By continuing to use the website, you consent to our use of cookies."

8 Data management

8.1 Data categories for terminological data collections

Depending on the purpose and goals of the TDC, different data categories may be required. For example, a terminological data collection for controlled authoring should contain usage notes that guide technical communicators in the correct use of, for instance, terms and proper names. A terminological data collection designed to support knowledge sharing should contain concept diagrams or concept models for users who absorb information more easily if it is presented in graphical form. It could also contain links to subject-specific ontologies.

Regardless of the purpose of the terminological data collection, each type of information recorded in a data field of a terminological entry shall be associated with a rigorously defined data category. Specifications for widely accepted data categories can be found at DatCatInfo[20], a data category repository that complies with ISO 12620.

Translation-oriented terminological data collections shall comply with ISO 30042 and shall follow the metamodel for terminological data collections described in ISO 16642 and ISO 30042:2019, Figure 2. The model is reproduced below in Figure 6.

```
                    ┌─────────────────────────┐
                    │           TDC           │
                    │(terminological data     │
                    │     collection)         │
                    └─────────────────────────┘
          ┌──────────────────┼──────────────────────┐
          ▼                  ▼                      ▼
┌──────────────────┐ ┌──────────────────┐ ┌──────────────────────┐
│        GI        │ │        CE        │ │         CI           │
│(global           │ │(concept entry)   │ │(complementary        │
│ information)     │ │                  │ │ information)         │
└──────────────────┘ └──────────────────┘ └──────────────────────┘
                              ▼
                     ┌──────────────────┐
                     │        LS        │
                     │(language section)│
                     └──────────────────┘
                              ▼
                     ┌──────────────────┐
                     │        TS        │
                     │ (term section)   │
                     └──────────────────┘
                              ▼
                     ┌──────────────────┐
                     │       TCS        │
                     │(term component   │
                     │    section)      │
                     └──────────────────┘
```
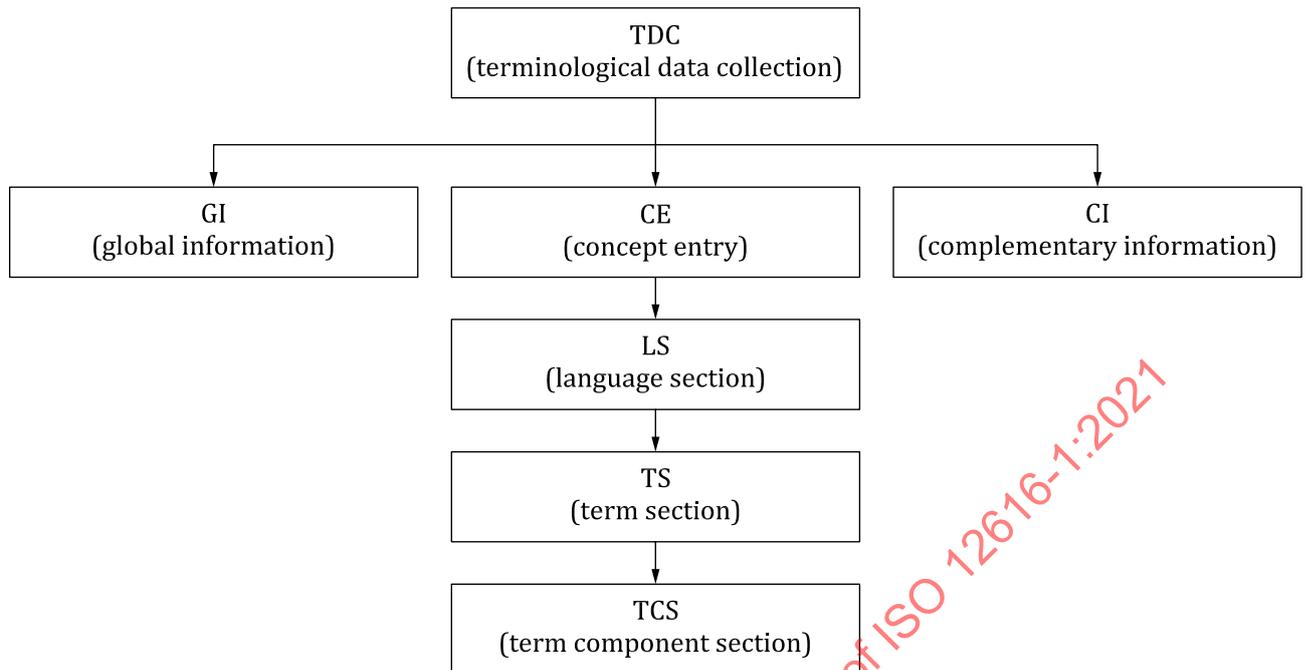
Source: ISO 30042:2019, Figure 2.

Figure 6 — Metamodel for terminological data collections

The three most important parts of the above model are "concept entry" (CE), "language section" (LS), and "term section" (TS). These are the levels where data categories are located. This document addresses data categories in terminological data collections designed to support translation environments. Annex A lists data categories that are typically used in translation-oriented TDCs. For more details see References [19] and [20].

The need for particular information in terminological entries differs. It depends on aspects, such as languages, the subject field, etc. Therefore, terminology workers should analyze which data categories are needed for their terminological data collections. Since the "term component section" (TCS) information container is typically not considered in translation-oriented terminography, it is not addressed in this document.

8.2   Principles of modelling concept entries

8.2.1   Guidance on designing terminological data collections

When building a terminological data collection, a terminology worker shall apply the principles of modelling concept entries as laid out in ISO 26162-1 and presented in 8.2.2 to 8.2.5, regardless of complexity of the TDC or of the tool used (see Clause 9). Examples of spreadsheets that meet these principles are provided in Annex B.

8.2.2   Concept orientation

In contrast to an entry in lexical resources, such as traditional dictionaries, a concept entry shall contain information about one concept.

EXAMPLE 1   The term "technical documentation" denotes two related, yet distinct concepts: (1) activity carried out by technical communicators, and (2) collection of technical documents pertaining to a specific product. Therefore, two concept entries shall be created, one for the process concept (1) and a second one for the results concept (2).

A subject field in one culture can develop differently from the same subject field in another culture. In such a case, the terminology worker must deal with differing concepts and concept systems, as concepts might differ in number, in their characteristics, and in the way they are differentiated from each other.

EXAMPLE 2   A source text refers to a legal concept which does not exist in the target language, due to differences in the two legal systems. There is, however, a concept in the target language that is very similar, fulfils the same main legal function, but differs on the level of characteristics.

If, for a given translation project, the differences between the concepts in question are so small that they can be neglected, the respective target-language designation can act as a so-called "quasi-equivalent". In such a case, the terminology worker may choose to document such quasi-equivalents in the concept entry, although, strictly speaking, source-language and target-language designations do not refer to the same concept. In this case, the terminology worker shall use a Transfer comment field or a Note field to document the quasi-equivalence of the designations and the degree of equivalence, as well as (where applicable) its directional nature (quasi-equivalence can work only for the given combination of source-language A and target-language B, but not in the reverse direction).

8.2.3   Term autonomy

A concept entry aims to provide an equal depth of information and description for each designation included in it. That does not mean that all data categories have to be filled in for all designations representing a concept. But the terminology worker shall have the option. In other words, the terminology database or spreadsheet shall be designed to offer an equal set of data categories and data category values for all designations in an entry.

EXAMPLE   The term "brake system" has been labelled as a "not recommended" synonym of the preferred term "braking system". Nevertheless, the same fields (Context, Part of speech, etc.) that are available for the term "braking system" are also available for the term "brake system".

8.2.4   Data elementarity

Individual data fields shall only hold one data element, i.e. a single unit of information that corresponds to the relevant data category and only one instance.

EXAMPLE   The Context field contains only a context (as explained in 6.4). It does not contain other information, such as the source of the given context. The source of the context is recorded in a Context source field. A Context field also contains only one actual instance of a sample text. If a second context is needed, a second Context field is used along with the corresponding Context source field.

8.2.5   Data granularity

Concept entries shall provide data categories in an appropriate level of detail.

EXAMPLE   When users of the TDC need detailed information about usage restrictions, it may be necessary to include fields for geographical usage, temporal restrictions, or proprietary information (trademarks, etc.). Otherwise, a simple Usage note field may be sufficient.

8.2.6   Repeatability

Since sometimes it is necessary to include more than one instance of a particular information type, such as multiple contexts as described in 8.2.4, Example, the corresponding field in the terminology database should be "repeatable", meaning that the user should be able to add another instance of the field in the entry. Further, it should be possible to include any field at more than one level of the entry (for example, a Definition field at the concept level and another at the language level).

EXAMPLE 1   A concept entry for translators usually contains designations in more than one language (repeatability by language).

EXAMPLE 2   A concept entry contains more than one designation for each language (repeatability within a language and by language).

EXAMPLE 3   Grammatical information (e.g. part of speech) is repeated for each designation (repeatability within a language).

EXAMPLE 4   A concept entry contains more than one context, possibly one for each designation. Each context, definition, etc. requires its own source information (repeatability within a language).

EXAMPLE 5   Figures and multimedia information may be stored in both the term section and the language section, depending on whether the need is to illustrate a universal concept or a locale-specific image, such as an image containing monolingual text (repeatability within a language and by language).

8.3   Open and closed data categories

Other data categories, such as in DatCatInfo[20], allow free text as their content. These are referred to as open data categories. When designing a field in a termbase that corresponds to an open data category, such as in DatCatInfo, the terminology worker shall allow the field to contain any text.

EXAMPLE 1   The data categories /term/, /definition/, and /context/ are typical open data categories.

Certain data categories, such as in DatCatInfo[20], require a limited set of values as their content. These are referred to as closed data categories, and the set of values is often referred to as a "picklist". Picklists have the advantage that the chosen values are consistent throughout the TDC.

EXAMPLE 2   In a picklist, one of the values for /part of speech/ is "noun". In such a database, the value in the Part of speech field for nouns would always be "noun", not a variation such as "Noun", "Common noun" or "common noun".

Other examples for data categories that are best set up as picklists are /grammatical gender/ and /term type/.

In terminology management systems, terminology workers can define pre-set values for specific data categories based on expectations about future content. This can save time in data input and helps to keep data consistent. Spreadsheets and other less sophisticated programs offer very limited ways to pre-set data categories, which is a distinct disadvantage.

When developing the data model, and specific types of information can be covered by a closed data category, preference should be given to closed data categories over open data categories, as this fosters data consistency in the TDC.

EXAMPLE 3   Since the number of subject fields relevant for a given project is limited, /subject field/ is defined as a closed data category.

When designing a field in a terminology database that corresponds to a closed data category, such as in DatCatInfo[20], the terminology worker shall restrict the content to a picklist.

8.4   Mandatory and optional data categories

Mandatory data categories constitute the information required for a concept entry to be complete and helpful for translators.

The following data categories shall constitute the minimum basic information required to create a concept entry for translators:

— /term/;

— /definition/ or /explanation/ or /context/;

— /part of speech/;

— /creation date/;

— /source/ or /source identifier/;

— /concept identifier/.

Optional data categories may be added to concept entries according to specific user needs. If, in a given project, the information needs of the intended users require that information of a specific type be recorded consistently, the respective data categories may be declared mandatory for this project.

8.5 Exchange and interoperability of terminological data

In low-complexity settings, the need for data exchange might not be immediately obvious (for some examples see 5.7). All TDCs shall support data exchange by complying with ISO 30042. For more details on TBX, see Reference [23].

9 Tools

There are various types of software tools that can be used for translation-oriented terminography. Table 3 below describes the most important types and the steps in the terminology management process where they can be applied. Terminology workers shall choose software tools that are best suited to the demands of a given project, the relevant process steps and application scenarios as detailed in in 5.3 through 5.8 and depicted in Figure 7.
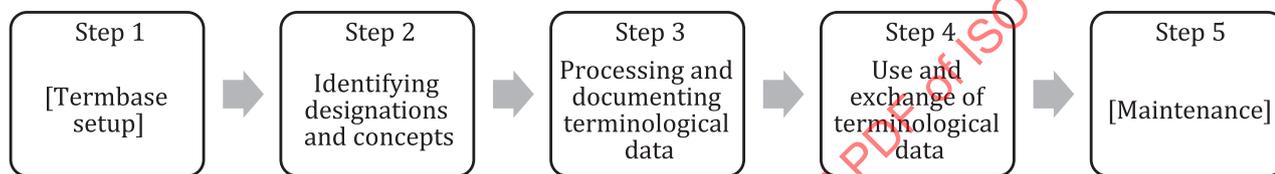

Figure 7 — Process steps

Table 3 — Software tools for translation-oriented terminography

Type of tool	Description	Step	Comments
text editor	enables a user to create and revise text (see ISO/IEC 2382:2015, 2126196)	Step 3	Text editors are not designed to manage terminology. However, terminology workers can use them to retrieve information, such as definitions or contexts.
spreadsheet program (also called spreadsheet)	displays a table of cells arranged in rows and columns, in which the change of the contents of one cell can cause recomputation of one or more cells based on user-defined relations among the cells (see ISO/IEC 2382:2015)	Steps 1–5	Since spreadsheet programs are not designed for terminographical purposes, terminology workers should use them only if no terminology management system (see below) is available.
text analysis tool (also called concordancer, concordancing software, corpus analysis software)	enables a user to examine language usage in a text or a text corpus, e.g. the frequency of terms, collocations, keyword-in-context lines, definitions	Steps 2–3	Some text analysis tools have similar features to term extraction tools (see below), but the latter are specifically designed for terminographical purposes.
term extraction tool	enables a user to identify and extract terminological data by searching through a text or a text corpus (see ISO 1087:2019, 3.5.6)	Step 2	Term extraction tools are specifically designed for terminographical purposes, although some text analysis tools (see above) have similar features.
terminology management system (TMS)	software tool with a metadata structure specifically designed for collecting, maintaining, and accessing terminological data (see ISO 1087:2019, 3.6.13)	Steps 1–5	TMSs can be stand-alone, cloud-based or client/server software tools.

10 Skills and competences

10.1 Terminology tasks and the required skills

Terminology workers establish or contribute to monolingual or multilingual terminological data collections. That work involves activities, such as identifying designations and the concepts represented by these designations, identifying characteristics of concepts, comparing concepts, ordering concepts in concept systems, describing concepts, and assessing or even creating designations. It also involves researching linguistic as well as subject-related information that is necessary for carrying out these activities. Last but not least, it involves processing and disseminating terminological data. For more information see Reference [17].

To carry out this work effectively, terminology workers shall possess specific skills and competences as discussed in 10.2 through 10.8. The terminology worker shall also have the following basic skills:

— broad general knowledge;

— project management skills;

— strong problem-solving skills;

— the ability to work systematically and accurately;

— an aptitude for abstract thinking, critical thinking; and

— creativity.

10.2 Basic terminology skills

Terminology workers shall possess knowledge of the fundamentals of terminology management and best practices in terminography, especially those covered in ISO 704. Terminology workers shall:

— comprehend the distinction between the concepts "concept", "object", and "designation" and be able to apply these concepts in terminology work;

— be aware of how characteristics form concepts and be able to identify characteristics expressed in texts or other media related to the concepts;

— comprehend the principle of concept orientation in terminology work and be able to apply it;

— comprehend the principles of drafting terminological definitions and be able to apply them to the assessment of existing definitions and the drafting of new definitions;

— comprehend how concepts can relate to each other, be able to identify hierarchical and associative concept relations, and know how to express these relations in terminological data collections and/or by means of graphical visualization;

— comprehend the concepts "synonymy", "homonymy", and "equivalence", be able to identify instances thereof in texts/media, and know how to deal with them in terminology work;

— know the major term formation methods applicable to the working languages;

— comprehend and be able to apply criteria for term assessment.

Terminology workers shall be mindful of the existence of other international standards on issues related to terminology and terminology work such as ISO 704, ISO 860, ISO 1087, and ISO 10241.

10.3 Information technology skills and competences

Terminology work involves using information and communication technologies (ICT) and typically the use of specialized terminology management systems. Terminology workers shall:

— have ICT skills, especially in working with office applications (e.g. spreadsheet or word processing programs), the Internet, and operating systems;

— be familiar with language-specific issues in the use of ICTs;

— know how to use search engines and specialized databases for the purpose of researching relevant information;

— comprehend the fundamentals of database management (data categories, entry structure, field types; see Clause 8);

— comprehend the fundamentals of terminological data modelling and be able to apply them in terminology work (see 8.2);

— be capable of working with specific terminology management systems or given alternatives (see Clause 9), in particular, to create and edit entries, to perform searches, to work with filters, and to perform standardized import and export procedures;

— be able to perform computer-aided term extraction;

— understand common data exchange problems and know how to solve them;

— be aware of relevant international and/or national standards (e.g. ISO 12620, ISO 30042);

— understand the basics of XML and be aware of related standards.

For recording and presenting terminological data, terminology workers should preferably work with terminology management systems that comply with international standards, such as ISO 16642, ISO 26162-1 and ISO 30042. If they do not have access to a terminology management system, they should store data in simpler formats, such as in spreadsheets or general databases, but make sure to follow guidelines (e.g. compatibility with export requirements or maintenance needs) to allow for use in a TMS in the future. In particular, terminology workers shall record and store terminological data in such basic IT applications in a way that implements the principles of modelling concept entries (see 8.2). Annex B provides examples on how to use spreadsheets to manage terminology.

10.4 Skills and competences related to term extraction

To find designations that should be included in a TDC, terminology workers can consult various relevant sources and select designations manually (e.g. by using a marker to highlight potential designations). Terminological data can also be extracted automatically with terminology extraction tools (see 5.4). To that end, terminology workers shall:

— have an excellent command of the languages involved;

— be able to assess a candidate term based on criteria, such as acceptability or clarity;

— be familiar with the term formation methods relevant in their working languages and in the respective subject field;

— be able to gauge whether a manual or automated approach is preferable for a given project;

— have a good understanding of existing approaches to computer-aided term extraction and their respective pros and cons;

— have a good understanding of the working mechanism used for term identification by the tool of choice and of the possibilities offered to enhance the result of the term extraction process.

10.5 Language skills

Terminology workers shall have thorough knowledge of their working language or languages, including:

— rules of grammar and orthography;

— language-specific methods of term formation;

— potential meanings of parts of words (such as stems, root words, prefixes, suffixes);

— linguistic requirements of terms (e.g. motivation, economy, correctness, naturalness), especially, when the terminology worker is involved in the assessment of existing terms or the coining of new terms (see ISO 704);

— differences in structure and grammar between working languages.

Terminology workers shall be able to apply this knowledge in term assessment (especially in the context of standardization processes and the formation of new terms).

Ideally, terminology workers should be first-language speakers of the language in which designations need to be coined. Knowledge of other languages is desirable to be able to compare formation patterns of designations. This is especially useful for languages where it is customary to avoid borrowings from another language.

10.6 Subject-matter expertise

A thorough knowledge of the subject matter at hand is a desirable skill, although it is not essential. If terminology workers do not have subject-matter expertise, they shall be aware of the importance of consulting subject-matter experts. This requires efficient networking and interpersonal skills and the ability to identify these specialists and persuade them to collaborate in terminology work.

10.7 Research competence

Terminology workers shall be capable of searching monolingual and multilingual resources for a given subject field in their relevant working languages. They need to be able to research information on concepts, especially definitions and other sources of information describing concepts, their characteristics, or their relation to other concepts. They also need to identify established designations as well as existing as-yet-non-standardized designations used by subject-matter experts. Furthermore, they need to research information on language use, such as term formation patterns, samples of typical usage, or collocations.

Terminology workers should know in what kind of resources (e.g. terminology standards, dictionaries, articles, books, glossaries, encyclopaedic resources, proprietary websites, corpora) the necessary information can be found and how to assess these resources (see 6.7). They should be able to distinguish between descriptive and prescriptive resources (see 4.4.3). Care should be taken to distinguish between high-quality materials and less reliable sources, and to identify locale-specific usage if possible.

10.8 Social and cultural competences

Terminology work is often a collaborative, knowledge-sharing effort which requires input from subject-matter experts. Thus, terminology workers need to interact effectively with people with different specialities and people from different cultures, organizations and companies. Often it is necessary to listen to others and encourage them to:

— share their views;

— clarify the reasoning behind strong opinions held by others (e.g. on a potential designation or definition);

— consider the views of others and empathize.