BS ISO 26162:2012

BSI Standards Publication

Systems to manage terminology, knowledge and content — Design, implementation and maintenance of terminology management systems

bsi.

...making excellence a habit.™

Document Sharing Hub

National foreword

This British Standard is the UK implementation of ISO 26162:2012.

The UK participation in its preparation was entrusted to Technical Committee TS/1, Terminology.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2012. Published by BSI Standards Limited 2012

ISBN 978 0 580 64664 5

ICS 01.020; 35.240.30

Compliance with a British Standard cannot confer immunity from legal obligations.

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 August 2012.

Amendments issued since publication

Date	Text affected

Documents
Document Sharing Hub

Full document details: BS ISO 26162:2012

INTERNATIONAL STANDARD

ISO 26162

First edition
2012-08-15

Systems to manage terminology, knowledge and content — Design, implementation and maintenance of terminology management systems

Systèmes de gestion de la terminologie, de la connaissance et du contenu — Conception, mise en œuvre et maintenance des systèmes de gestion de la terminologie

Reference number
ISO 26162:2012(E)

© ISO 2012

Documents
Document Sharing Hub

Full document details: BS ISO 26162:2012

COPYRIGHT PROTECTED DOCUMENT

© ISO 2012

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

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub

Full document details: BS ISO 26162:2012

Contents

Page

Foreword vi

Introduction vii

1 Scope 1

2 Normative references 1

3 Terms and definitions 1
3.1 Resources 1
3.2 Data categories 3
3.3 Data modelling 5
3.4 Applications 7

4 Terminology management system (TMS) 7
4.1 General description 7
4.2 Purchased or self-programmed TMS 8
4.3 Predefined or freely definable TMS 8
4.4 Desktop, client-server or Web-based TMS 8
4.5 Stand-alone, integrated or combined TMS 8
4.6 Monolingual, bilingual or multilingual TMS 9
4.7 Database or structured text TMS 9
4.8 Single-database or multiple-database TMS 9

5 Project phases 9
5.1 Overview 9
5.2 Pre-feasibility study 10
5.3 Feasibility study 10
5.4 Use-case analysis 10
5.5 System requirements 10
5.6 Cost-effectiveness 11
5.7 System design 11
5.8 System development 11
5.9 System deployment 11
5.10 System test 12
5.11 TMS population, use, and maintenance 12

6 User-centred design 12
6.1 Basic procedures 12
6.2 Steps in a user-centred approach 13
6.3 Identifying the users and their needs 13
6.4 Identifying the output products 16
6.5 Performing a task analysis and preparing use cases 16
6.6 Identifying and prioritizing requirements 18
6.7 Conducting a competitive evaluation 19
6.8 Designing and evaluating a prototype 19
6.9 Adjusting the design to user feedback 19
6.10 Performing a beta assessment 19

7 Terminological data categories 19
7.1 Introduction to data categories 19
7.2 Principles for selecting and using data categories 20
7.3 Types of data categories 23
7.4 Data entry structures 26
7.5 Selecting data categories 26
7.6 Translation-specific data categories 27

Documents
Document Sharing Hub

© ISO 2012 – All rights reserved

Full document details: BS ISO 26162:2012

7.7 Prescriptive data categories ... 27
7.8 Workflow-related data categories ... 27
7.9 Standardized data category names and data category concepts ... 27

8 Data modelling ... 28
8.1 The terminological metamodel ... 28
8.2 Data modelling for concept orientation ... 29
8.3 Application-oriented approaches ... 32
8.4 Data modelling examples ... 33
8.5 Accounting for legacy data ... 39

9 Implementing the TMS ... 40

10 Deploying the TMS ... 41
10.1 Deployment activities ... 41
10.2 Preparing documentation, help, and educational materials ... 41
10.3 Providing support and maintenance ... 42
10.4 Meeting stakeholder dependencies ... 42
10.5 Announcing and promoting the TMS ... 42
10.6 Delivering the TMS ... 42
10.7 Providing training ... 42
10.8 Measuring user satisfaction ... 42

11 User interfaces ... 43
11.1 Designing the user interface ... 43
11.2 Displaying terminological data categories ... 43
11.3 Displaying and arranging terminological entries ... 44

12 Inputting and editing data ... 47
12.1 Manually inputting data ... 47
12.2 Importing data ... 51
12.3 Editing data ... 52
12.4 Validating data ... 52
12.5 Automatically generating or modifying data ... 53
12.6 Adding cross-references ... 53
12.7 Adding multimedia files ... 53

13 Search functions ... 54
13.1 Database search features ... 54
13.2 Searching for a term ... 55
13.3 Searching by concept number or characteristics ... 55
13.4 Complex filtering and searching ... 56
13.5 Searching in text fields ... 56
13.6 Browsing ... 56

14 Data output ... 56
14.1 Types of data output ... 56
14.2 Displaying search results ... 56
14.3 Sorting ... 57
14.4 Printouts ... 58
14.5 Exporting the data to a file ... 61
14.6 Exporting data for other applications ... 61

15 Organizing and administering the TMS ... 61
15.1 Creating a management plan ... 61
15.2 Importance of data flow management ... 62
15.3 Changing the data model ... 62
15.4 Providing for data security ... 62
15.5 Controlling access ... 63
15.6 Supporting an interchange format ... 63
15.7 Staffing the TMS ... 64
15.8 Controlling costs and managing resources ... 64

Annex A (informative) **Case studies: Data categories and data modelling** ... 66

Documents
Document Sharing Hub

© ISO 2012 – All rights reserved

Full document details: BS ISO 26162:2012

Bibliography **69**

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

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 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 26162 was prepared by Technical Committee ISO/TC 37, *Terminology and other language and content resources*, Subcommittee SC 3, *Systems to manage terminology, knowledge and content*.

© ISO 2012 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

Introduction

Terminological data are collected, managed and stored in a wide variety of terminology management systems (TMSs). TMSs employ a variety of database management systems, ranging from personal computer applications for individual users to client server applications or Web-based applications operated by major companies and governmental agencies. Terminological data collections (TDCs) are based on various kinds of data models and consist of different sets of data categories (Data Category Selections, DCSs). To facilitate co-operation and to prevent duplication of work, it is important to develop standards and guidelines for creating and using TDCs as well as for sharing and exchanging data.

ISO/TC 37 has published the following standards in order to facilitate the exchange of terminological data and to create an integrated approach to be used in analysing an existing TDC and in designing new ones: ISO 704, ISO 12620, ISO 16642.

Having some acquaintance with these documents before reading this International Standard would be beneficial.

This International Standard gives guidance on choosing the relevant data categories, as well as designing and implementing a data model and a user interface for a TMS, taking into account the intended user group. The guidelines described in this International Standard are indispensable for the successful development of a TMS and for avoiding costly errors. This International Standard may be used for choosing the appropriate TMS for a certain purpose.

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

Documents
Document Sharing Hub

INTERNATIONAL STANDARD ISO 26162:2012(E)

Systems to manage terminology, knowledge and content — Design, implementation and maintenance of terminology management systems

1 Scope

This International Standard specifies criteria for designing, implementing and maintaining terminology management systems (TMSs).

This International Standard provides information about the rationale for using a TMS, types of users and users' needs, steps in designing and implementing a TMS, as well as the tasks of organizing and managing a terminological data collection (TDC). It also provides guidelines for selecting and using data categories for managing terminology in various environments.

This International Standard is intended for terminologists, software developers and others who are involved in the process of developing or acquiring a TMS.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 704, *Terminology work — Principles and methods*

ISO 12620, *Terminology and other language and content resources — Specification of data categories and management of a Data Category Registry for language resources*

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

ISO 30042, *Systems to manage terminology, knowledge and content — TermBase eXchange (TBX)*

3 Terms and definitions

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

3.1 Resources

3.1.1
terminological resource
terminological data collection
TDC
text or data resource consisting of **terminological entries** (3.1.4)

NOTE Adapted from ISO 24613:2008.

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

3.1.2
terminology management system
TMS
software tool specifically designed for collecting, maintaining, and accessing terminological data

3.1.3
terminological database
TDB
termbase
database comprising a **terminological resource** (3.1.1)

3.1.4
terminological entry
TE
part of a **terminological resource** (3.1.1) that contains the terminological data related to one concept

NOTE Adapted from ISO 1087-2, 2.22. The abbreviated term has been added.

3.1.5
concept orientation
principle applied to terminology management whereby a **terminological entry** (3.1.4) describes one and only one concept or two or more **quasi-equivalent concepts** (3.1.7)

NOTE A concept-oriented entry can contain multiple terms, which are therefore considered semantically equivalent.

3.1.6
equivalent concept
concept in one language that comprises the same characteristics as that covered by a given concept in another language

3.1.7
quasi-equivalent concept
nearly equivalent concept
concept in one language that shares most but not all characteristics with a concept in another language, but that is nevertheless used as an equivalent for that concept in some contexts

3.1.8
entailed term
term used in a text field such as the */definition/* or */context/* that designates a concept that is defined in another **terminological entry** (3.1.4) in the same **terminological resource** (3.1.1)

3.1.9
doublette
terminological entry (3.1.4) that describes the same concept as another entry

NOTE Doublettes are normally detected by TMS by identifying two terms having the same form; however, doublettes are determined based on identical concepts, not on identical terms. Doublettes should not be confused with homographs.

3.1.10
concept system
set of concepts structured according to the relationships among them

[ISO 1087-1:2000, 3.2.11]

3.1.11
concept diagram
graphic representation of a **concept system** (3.1.10)

[ISO 1087-1:2000, 3.2.12]

© ISO 2012 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

3.1.12
legacy data
terminological data that are available in an existing file or database, and that are considered for importation into a **TMS** (3.1.2)

NOTE Legacy data may be in the form of previously used databases, word-processing files, comma-delimited text files, SGML, HTML, or XML files, and the like. Conversion of such data to a format that will be compatible with a new TMS can pose serious challenges.

3.1.13
term
word, or several words, that denote a concept

EXAMPLES "Olympics" and "special Olympics" are two terms in the field of sports.

NOTE 1 When the word or words can denote more than one concept, each word/concept pair is a separate term. For example, "port" (shelter for boats) and "port" (computer connection point) are two different terms.

NOTE 2 In terminology theory, terms denote concepts in specific subject fields, and words from the general lexicon are not considered to be terms. In a TDC, however, words from the general lexicon are sometimes recorded in terminological entries, where they are still referred to as "terms".

3.2 Data categories

3.2.1
data category
result of the specification of a data field

[ISO 1087-2:2000, 6.14]

3.2.2
data element
unit of data that, in a certain context, is considered indivisible

[ISO 1087-2:2000, 6.11]

3.2.3
data granularity
degree of precision of data

NOTE For example, the set of individual **data categories** (3.2.1) */part of speech/*, */grammatical gender/,* and */grammatical number/* provides for greater data granularity than does the single data category */grammar/.*

3.2.4
data elementarity
principle whereby a single data field shall contain only one item of information

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

3.2.5
term autonomy
principle whereby all terms in a **terminological entry** (3.1.4) can be described by using the same set of **data categories** (3.2.1)

3.2.6
Data Category Registry
DCR
set of standardized **data categories** (3.2.1) to be used as a reference for the definition of linguistic annotation schemes or any other formats in the domain of language resources

[ISO 12620:2009, 3.2.1]

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

NOTE The ISO/TC 37 DCR contains **data category specifications** (3.2.7), which include historical, descriptive, and administrative information and other metadata.

3.2.7
data category specification
set of attributes used to fully describe a given **data category** (3.2.1)

[ISO 12620:2009, 3.2.2]

NOTE The abbreviation DCS refers to **Data Category Selection** (3.2.8).

3.2.8
Data Category Selection
DCS
set of **data categories** (3.2.1) selected from the **Data Category Registry** (3.2.6)

[ISO 12620:2009, 3.2.3]

3.2.9
complex data category
data category (3.2.1) that has a **conceptual domain** (3.2.11)

[ISO 12620:2009, 3.1.7]

3.2.10
open data category
complex data category (3.2.9) whose **conceptual domain** (3.2.11) is not restricted to an enumerated set of values

[ISO 12620:2009, 3.1.8]

3.2.11
conceptual domain
set of valid **value meanings** (3.2.14)

NOTE 1 Based on ISO/IEC 11179-1:2004, 3.3.6.

NOTE 2 The value meanings in a conceptual domain may be enumerated, further specified by additional constraints or expressed via a description. For instance, the **data category** (3.2.1) */term/* is described by its definition and thus constrained from properly containing, for example, contextual information or grammatical information, but it would be impossible to enumerate all values associated with this data category.

3.2.12
closed data category
complex data category (3.2.9) whose **conceptual domain** (3.2.11) is restricted to a set of identified **simple data categories** (3.2.13)

[ISO 12620:2009, 3.1.13]

3.2.13
simple data category
data category (3.2.1) with no **conceptual domain** (3.2.11)

[ISO 12620:2009, 3.1.12]

3.2.14
value meaning
meaning or semantic content of a value

[ISO/IEC 11179-1:2004, 3.3.39]

© ISO 2012 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

NOTE ISO/TC 37 treats enumerated values as simple data categories, that is, as data categories in their own right. The meaning of a value is always viewed in the context of the overall value domain and the closed data category with which it is associated, and is not just a property of the domain value itself.

3.2.15
value domain
set of **permissible values** (3.2.16)

[ISO/IEC 11179-1:2004, 3.3.38]

3.2.16
permissible value
expression of a **value meaning** (3.2.14) allowed in a specific **value domain** (3.2.15)

[ISO/IEC 11179-1:2004, 3.3.28]

3.2.17
thematic domain
class of applications identified by the similarity of the data structures they need to manipulate

[ISO 12620:2009, 3.4.3]

EXAMPLES Terminology, lexicography, morphosyntactic annotation.

3.2.18
thematic domain profile
profile
representation within a **data category specification** (3.2.7) of the **thematic domain** (3.2.17) with which a **data category** (3.2.1) is associated

[ISO 12620:2009, 3.4.4]

NOTE A data category may have several thematic domain profiles, indicating that it is used by several thematic domains.

3.3 Data modelling

3.3.1
data model
graphical and/or lexical representation of data, specifying their properties, structure and inter-relationships

[ISO/IEC 11179-1:2004, 3.2.7]

3.3.2
data modelling
process of structuring and organizing data, typically for implementation in a database management system

3.3.3
data modelling variance
variation in the assignment of **data categories** (3.2.1) to data models as a result of differences in philosophy with respect to the ordering of information in the **terminological entry** (3.1.4)

3.3.4
metamodel
data model (3.3.1) that specifies one or more other data models

[ISO/IEC 11179-1:2004, 3.2.20]

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

3.3.5
metadata
data that define and describe other data

[ISO/IEC 11179-1:2004, 3.2.16]

3.3.6
global information
GI
technical and administrative information applying to a complete data collection

[ISO 16642:2003, 3.7]

EXAMPLE Title of the data collection, revision history.

3.3.7
complementary information
CI
information supplementary to that described in **terminological entries** (3.1.4) and shared across the **terminological data collection** (3.1.1)

[ISO 16642:2003, 3.1]

EXAMPLE Domain hierarchies, institution descriptions and bibliographical references are typical examples of complementary information.

3.3.8
shared resource
information object that can be accessed from any of the terminological or lexicographical entries in a terminological or lexicographical resource

EXAMPLE Shared resources typically include bibliographical entries, responsibility entries, namespace identifiers, frequently referenced textual material, geographical location lists, and external files, such as graphics or audio files.

3.3.9
language section
LS
part of a **terminological entry** (3.1.4) containing information related to one language

[ISO 16642:2003, 3.9]

3.3.10
term section
TS
part of a **language section** (3.3.9) containing information about a **term** (3.1.13)

[ISO 16642:2003, 3.15]

3.3.11
class
object class
⟨UML⟩ description of a set of objects that share the same members

3.3.12
multiplicity
number of instances of one **class** (3.3.11) linked to one instance of another class in a relationship set

NOTE For instance, an origination date may be the same for many terminological entries, but one terminological entry can only have one origination date.

© ISO 2012 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

3.4 Applications

3.4.1
language planning
deliberate efforts to influence human behaviour with respect to the acquisition, structure, or functional allocation of language

EXAMPLE Language planning can include the standardization of spelling and grammatical rules, the specification of official national languages, efforts to create viable neologisms designed to increase the capacity of a language as a vehicle for scientific and commercial communication, and measures designed to protect the language from foreign influences.

3.4.2
descriptive terminology
approach for managing terminology that documents the way that **terms** (3.1.13) are used in contexts without indicating preferred usage

3.4.3
prescriptive terminology
approach for managing terminology that indicates preferred usage

3.4.4
normative terminology
approach for managing terminology that is used in standards work or governmental regulation

3.4.5
translation editor
software that supports the process of creating and revising translations

3.4.6
controlled authoring
authoring that uses limited vocabulary and textual complexity to produce clear documents

3.4.7
localization
l10n
process of taking a product and making it linguistically and culturally appropriate to the target locale (country/region and language) where it will be used and sold

[Localization Industry Standards Association]

4 Terminology management system (TMS)

4.1 General description

A TMS is a software tool specifically designed to collect, maintain, and access terminological data. It is used by translators, terminologists, technical writers, and various other users. TMSs shall be based on the principles for terminology work set down in ISO 704, which means that the terminological resources that are created by using a TMS contain terminological entries designed to document concepts and all the terms (synonyms and equivalents in other languages) that denote those concepts.

TMSs are typically used to elaborate, maintain and disseminate TDCs in a variety of organizational environments, including national bodies, government agencies, interest groups, companies, and individuals. Depending on the organizational framework, the TMS will need to support a variety of necessary features, different tasks and potential users.

When planning a software application for terminology management, it is important to consider the different possible types or categories of TMS.

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

4.2 Purchased or self-programmed TMS

There is a variety of terminology management tools on the software market; most of the software is commercial, but an increasing array of freeware or open-source software is available for terminology management. If the organizational, functional and methodological needs of the respective application environment are met by the system, it can be used off the shelf. However, in many cases, a customer-specific adaptation and adjustment is necessary. If the commercial software cannot be tailored to specific, identified needs, a TMS will have to be developed in-house. Costs, maintenance, support and interoperability are important parameters to be considered when deciding between buying or leasing a TMS, and developing one in-house.

4.3 Predefined or freely definable TMS

Commercial TMSs can have predefined or freely definable data modelling structures. A TMS with a fixed set of data categories, a defined set of languages and/or an unchangeable data model shall only be used if the setting corresponds exactly with the needs of the organizational environment in which the TMS will be applied. A TMS with a freely definable structure allows users to define their own data categories and their own entry structures so that the software can be adapted to suit users' specific terminological needs and can be modified as future requirements change.

4.4 Desktop, client-server or Web-based TMS

The simplest TMSs are installed on a local computer for a single individual. Sometimes the underlying database files are stored on a LAN computer and can be shared by multiple desktop clients. In large organizations with many users (and different roles), only a TMS with client-server architecture can serve the needs of the organization. These TMSs require a client installation on each local desktop and a server software installation handling the simultaneous access of different clients (with different tasks).

If multiple users all over the world are to have access to the terminological data, the only viable solution is a Web-based TMS; users do not have to install software specific to the TMS on their local computers because a standard Web browser will allow all operations with the TMS. When opting for a Web-based TMS, designers also have to decide whether to host their own server or to utilize a third-party TMS, for instance choosing between services offered by major TMS software suppliers or free, Web-resident collaborative environments. Before opting to outsource the hosting function, they should carefully weigh factors such as cost, host expertise, and reliable data accessibility, as well as data security and confidentiality.

4.5 Stand-alone, integrated or combined TMS

A stand-alone TMS is an autonomous software package designed for managing terminology independently of any other application. Integrating the terminology from a stand-alone TMS into other applications, such as controlled authoring tools or Computer Assisted Translation (CAT) tools, usually requires terminology to be exported from the TMS and imported into the other application. Sometimes the integration can be realized by developing specific plug-ins for each application in which the terminology is required. Because they are not tied to a specific software application, these types of TMS typically offer more flexibility in terms of the adaptability of the data for various applications.

An integrated TMS is one that comprises a set of terminology management functions directly within another software application, such as in a CAT tool. In order to use an integrated TMS, the user has to install and use the TMS package in conjunction with a seamlessly interoperable suite of software. These types of TMS are designed to offer the functions and features specifically required by the software of which they are a part, but, as a consequence, they might neglect to include features required for other applications. For instance, a TMS integrated into a controlled authoring application will contain functions and data types required for controlled authoring, but it would probably lack features required for the translation process.

© ISO 2012 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

4.6 Monolingual, bilingual or multilingual TMS

Monolingual TMSs are generally used by authors and readers who are probably not involved in translation or multilingual document production. A bilingual TMS can feature simple term or word pairs, sometimes using a lexicographical approach, where Term A is represented as equal to Term B, without reference to any subject field, definition or context. They may also respect concept orientation, where the terms listed in the two languages pertain specifically to a single concept, and if a term is attributed to more than one concept, then each concept is recorded in a separate entry. A multilingual TMS is usually concept-oriented because of the difficulties involved in managing multiple meanings across multiple languages; some of these TMSs also allow for a bilingual view, for instance for use in translation environments.

4.7 Database or structured text TMS

Since a TMS has to maintain a huge amount of data, a database engine is usually running in the background. A relational database is the ideal architecture for handling typical terminological requirements such as many terms in many languages, multiple data categories, and (unlimited) repeatability of specific types of terminological information. A structured text-oriented TMS also meets many of the typical terminological requirements by tagging the different types of data, such as in XML.

4.8 Single-database or multiple-database TMS

A TMS can store all terminological entries in one single database. Different subject fields, processing stages, customers and other types of information can be classified by attributes, and filters can be used to create a virtual subset of the database. Other TMSs store the terminological data in different physical databases, such as for different subject fields or different customers. Either the user has to select one of the databases before searching and editing, or the software allows for a simultaneous search and access to all databases, sometimes with a definable hierarchy specifying which database will be consulted first. A special case of the latter is a federated TMS that collects the data from different external databases, potentially with different data structures.

5 Project phases

5.1 Overview

As described in 4.2, for some purposes it may be possible to buy or otherwise utilize a TMS; however, in other cases, it may be necessary to develop a new TMS. The process of implementing a TMS, whether it is a third-party or a proprietary solution, comprises some or all of the phases listed below. Some projects, especially agile development projects which may also involve end users, require an iterative process where several phases are repeated until the desired outcome is achieved.

— Conduct a pre-feasibility study (see 5.2).

— Conduct a feasibility study (see 5.3).

— Conduct a use-case analysis (see 5.4).

— Establish system requirements (see 5.5).

— Analyse cost-effectiveness (see 5.6).

— Design the TMS (see 5.7).

— Develop the TMS (see 5.8).

— Deploy the TMS (see 5.9).

— Test the TMS (see 5.10).

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

— Populate, use and maintain the TMS (see 5.11).

A detailed work and time schedule should be established to set deadlines and distribute the tasks to the personnel involved. Intermediate reports should document the results. The schedule should clearly indicate when decisions will be made to proceed with the project (so-called go/no-go decision) and whether to build a new TMS, buy or opt for an existing TMS, or modify an existing TMS. These decisions may recur at different stages in the process as the requirements are prioritized.

The phases described in this clause can be adjusted according to the scope of the project and the type of organization involved (commercial enterprises, government agencies, educational institutions, translation agencies, and so forth).

5.2 Pre-feasibility study

The pre-feasibility study defines the scope of the project and determines which user groups, organizational units and systems should be involved. Stakeholders and users can include people with specific roles in the organization (writers, translators, and so forth), organizational entities such as publishing departments and translation services, business partners and vendors, and customers. But they can also include persons responsible for related software and systems such as file handling systems, workflow systems, translation software, and content authoring tools.

There may be a decision to deliberately exclude units or parts of the process defined here or to postpone them to a later stage of the project. The pre-feasibility study also provides an initial description of the process of creating the TMS. It justifies the need for introducing a TMS, including identifying any alternatives, and determines the basic requirements of the potential users. The organizational environment required to implement and operate the TMS should be discussed.

5.3 Feasibility study

A feasibility study includes a more refined outline of the tasks to be performed and the personnel and financial resources necessary to establish and operate the TMS. The feasibility study should also consider the technical and organizational integration of the TMS into the overall information management environment of the organization. The feasibility study should enable a final decision to be made on whether or not the TMS shall be implemented.

5.4 Use-case analysis

Conducting a use-case analysis is essential for ensuring that the TMS will meet the needs of the target user group. This phase is described in 6.3.

5.5 System requirements

During this phase, the specific requirements for the projected TMS are identified based on feedback from direct users and other stakeholders. Software feature specifications are then defined to fulfil those requirements. The hardware and software components are chosen according to the prerequisites outlined in the software feature specifications.

Typically, requirements are prioritized based on user feedback since it may not be possible to deliver all requirements (see 6.6).

In addition to user-centred requirements, technical requirements relating to the software infrastructure of the organization shall be met, such as to ensure conformity to company standards, as well as security and maintenance procedures.

The data model for the TMS shall be defined before it is developed or purchased.

© ISO 2012 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

5.6 Cost-effectiveness

The costs of implementing and operating the TMS should be estimated, as well as the expected benefits of the system. It should be noted that the financial return on investment will not be realized until after the TMS has been operating for a while. Any business case that is used to determine whether or not the project is approved for funding should include intangible, non-financial benefits (sometimes called “soft benefits”) in addition to financial ones. Soft benefits to consider include improved quality and customer satisfaction, cost avoidance through reduction of work duplication or other sources of inefficiencies, and more efficient processes (such as translation and customer service). If several alternative options are still possible at this phase (such as building a new TMS or buying a commercially available TMS), the costs and benefits should be estimated and compared for each option.

Another factor to consider is the cost of *not* implementing a TMS. The cost-avoidance factors mentioned above fit into this category, but there are other considerations such as the legal costs of trademark infringement that a TMS might help to avoid, the lost opportunities for controlled authoring and the additional difficulties in implementing a content management system without controlled terminology. All the potential current and future applications of terminology should be taken into account in building the business case.

The costs of importing existing data can be significant, particularly if the existing data set needs to be modified prior to import. The facilities for importing the existing data offered by the various systems being considered should be calculated in the cost analysis. It may be possible to request that the supplier of an existing TMS develop any required custom conversion routines.

5.7 System design

This stage involves either searching for and finding software products that can be used to implement the TMS, or designing a new TMS. If several software products are available on the market that may meet the requirements, each will have to be evaluated according to the projected requirements for the TMS – including the cost factor – in order to determine the best solution. If a new TMS is to be created, a detailed concept for the software system and its components has to be developed and the environment and staff for developing the software has to be provided.

Any requirements for importing existing terminological data shall be considered in the design.

5.8 System development

The objective of this stage is to develop the necessary hardware and software components. The TMS and its subsystems are implemented by writing new software or by configuring and/or customizing existing software. It is essential to document and test the components and their function within the whole system as they are implemented.

5.9 System deployment

The deployment stage encompasses several substages and should ideally be integrated with the testing stage.

Initial deployment, especially in large enterprises, should feature a system launch for a small group of selected users. During this initial launch, the TMS should be tested and any shortcomings should be corrected. The individuals involved in this initial deployment should participate in testing activities and should experiment with a small set of terminological entries before working with large projects.

Full enterprise-wide deployment should be undertaken only after completion of the testing phase, as described in 5.10.

© ISO 2012 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

5.10 System test

A thorough test of the whole system (all the software features and requirements) shall be performed after initial deployment. Testing should be based on the complete processing of a representative set of terminological entries that is large enough to challenge the system. The test data should be discardable in case of testing failures. Testing shall cover all aspects of terminology workflow and include multiple-use case scenarios. Testing should encompass all program features (input, output, import, export, displays, etc.) and involve representatives of all user roles (terminologists, technical writers, subject field specialists, translators, etc.). Manual testing can be enhanced by the use of automatic methods and routines and can provide initial data that can be used later to analyse the benefits of changes in the system.

Individuals should not be allowed to set up their own stand-alone environment of the same TMS until the testing is complete.

5.11 TMS population, use, and maintenance

As soon as the TMS has been deployed and tested, the system is operational. Full deployment involves instructing system administrators and system users, establishing the required maintenance infrastructure, and integrating the TMS into the overall workflow. It may be necessary to populate the TMS initially by importing existing terminological data. Prior to import, the data should be reviewed; modifications in content, structure, and format may be necessary. This task can take considerable time so the effort should be estimated and appropriate time allowed in the overall schedule. For more information about importing data, see 12.2.

6 User-centred design

6.1 Basic procedures

6.1.1 User-centred design refers to the methodologies and principles of designing a product so that it meets the needs of the product's users. The purpose of this clause is to explain how to include the users in the design process.

The potential users of a TMS include:

— people, such as terminologists, translators, interpreters, subject field specialists, teachers, students, technical writers, advertisers, media designers and scriptwriters, knowledge and information engineers, language engineers, product developers, engineers, and so forth;

— machines and tools that access and process terminological information, such as controlled authoring tools, spelling and grammar checkers, machine translation engines, term extraction tools, computer-assisted translation tools, concordance functions, search engines, text analysers, and so forth.

Every user category has its particular characteristics and specific needs. If the TMS is intended for more than one user category, the common denominator or dominant needs of all users should be determined, and these needs should be prioritized over all others.

6.1.2 Designers should determine:

— whether the goal of the terminology collection is normative, prescriptive, or descriptive;

— whether the TMS should be a general one covering all fields or a specialized one treating only one subject field or a limited number of subject fields and, if it is the latter, which subject fields shall be covered.

6.1.3 Options for delivering information to users with varying needs include providing:

— multiple views of the terminological data (partial views, some fields hidden until requested);

© ISO 2012 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

— all information available in all languages, or some information available in only one language;

— different interfaces for various purposes (translator's workstation, technical writing support, interaction with other computerized tools, etc.);

— access on a local (in-house), national, and/or international basis;

— Web interface or another type of interface;

— open public access, or access restricted to a certain environment.

6.1.4 Designers shall determine how the database is going to be used. Some types of information are more important to some users than to others. For example, technical writers need to know the meaning of technical terms and how to write the term correctly (capitalization, hyphenation, etc.), as well as what term to choose when they are aware of synonyms. Translators are typically interested in finding the preferred equivalent for a source term in a given target language. Software programs that interact with the TMS, such as those mentioned above, should be able to access the different types of information they require.

6.2 Steps in a user-centred approach

Adopting a user-centred approach to designing a TMS helps to ensure that the system will be easy to use and will achieve the desired objectives.

The following key steps are described.

— Identify the users and their needs (see 6.3).

— Identify the output products (see 6.4).

— Analyse the tasks to be performed and create use cases (see 6.5).

— List the requirements (see 6.6).

— Conduct a competitive evaluation (see 6.7).

— Prepare and evaluate a design prototype (see 6.8).

— Adjust the design according to user feedback (see 6.9).

— Perform a beta assessment (see 6.10).

6.3 Identifying the users and their needs

The first step in designing a TMS is to determine who will use it and what their needs are. All potential users of the TMS should be surveyed to find out if they will use it, how often they will use it, what tasks they will perform, and what specific needs they have. Designers can also conduct focus group sessions and one-on-one interviews to probe more deeply into the responses. The TMS should be designed to meet these needs, and should be assessed at several stages during its development to determine how successfully it meets these needs and expectations.

The different types of users should be surveyed in a manner that is proportional to their use. For example, if there are 1 000 translators and 100 terminologists who are expected to access the TMS on the Web, then the ratio of translators to terminologists who are surveyed on the Web-based functions should be 10:1. However, if the terminologists will spend twice as much time updating the data collection as translators, the ratio should be adjusted accordingly for questions about the update functions: 10*1:1*2 = 10:2.

Special attention should be given to needs that seem to contradict each other. Sometimes, the perceived needs of one group are opposed to the needs of another group; for instance, technical writers and translators

Documents
Document Sharing Hub

© ISO 2012 – All rights reserved

Full document details: BS ISO 26162:2012

can express mutually exclusive requirements. For example, context sentences (sample sentences containing the term) may be useful for translators but not for writers. In such cases, it may be possible to meet both needs by providing user-specific views of the data, and/or user-specific workflows.

Online surveys can be used to collect feedback from off-site users who cannot attend a face-to-face focus group session. The preparation of a survey should include a pilot test of the survey, in which the survey is tested with a small group, before conducting the survey with the complete user group. This will help ensure that the right questions are asked in the right order. The focus group for the pilot test should be representative of all types of users in order to avoid biased conclusions, and steps should be taken to motivate users to participate in the survey.

To validate needs and the perception that survey-takers have of the needs they express, questions should be included that attempt to evaluate the results that the survey-takers envision will be obtained by implementing their suggestions. It is important to avoid any ambiguity in survey questions.

In addition to identifying the needs of the various user groups, the survey should provide information about the following general determining factors, taking into consideration both present and potential future needs.

1. What is the main purpose of the TMS?

   — For example, if the goal is to *standardize* terminology in one or more languages, then the data collection should be *normative*. This means that it will enforce the use of specific standard terms. In this case, the design and methodology adopted for the TMS should follow the ISO standards for standardizing terminology, such as ISO 704, ISO 1087, and ISO 10241-1. On the other hand, if the goal is to *improve* the use of terminology, then the data collection should be *prescriptive*. This means that it should indicate preferred terms and non-preferred terms and include usage guidelines. If the goal is to *describe* how terms are used (without prescribing how they should be used), then the data collection is *descriptive*. This type of terminological resource does not indicate preferred or non-preferred terms.

   — If the data collection is meant to support a localization service provider that has many different clients, then the integrity of the client's terminology within the TMS is paramount. Data fields should be provided to identify the origin of terms, the client, and the projects where the terms are used, and security measures may be implemented allowing for the protection of the data provided by, or developed for, the client. In some cases, separate data collections should be provided for separate clients.

2. Are all users within the company/organization or are some external?

   — The answer to this question will impact the delivery and access channels for the data. For example, in the case of a network-based TMS, if some users are outside the organization's firewall, special measures should be taken to provide them with access to the TMS. Will all the terminology data be suitable for both internal and external users? If not, it will be necessary to somehow differentiate data that can be distributed externally from data that cannot, such as through the use of a special data field. External groups such as customers and vendors may not have access to the same tools for using the terminology. In this case interoperability of the data is important and shall be facilitated through the support of international interchange standards (see ISO 12620, ISO 16642, and ISO 30042).

   — Whenever multiple versions or copies of the data exist internally and externally, versioning and global updating should be accounted for in workflow and scheduling. It is, however, not advisable to maintain multiple versions of data.

3. Are all users experienced language professionals (writers, translators, terminologists, etc.) or are some from other areas such as sales, management, support services, training, and so forth?

   — A TMS that serves a range of user types may need to display different types of information to different types of users. If many users can create terminological entries, they also may need different interfaces to create those entries. Administrative personnel will not want to record as much detailed

© ISO 2012 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 26162:2012

information as terminologists, and technical writers may need to have information presented in different ways than that required by translators.

4. Will any other tools, such as a desktop publishing system or a translation memory system, need to access or interact with the TMS?

   — The design phase should identify all tools that may interact with terminological data and determine what data and formats they require. Machine processes can require data that are not typically required by human users, such as the part of speech.

5. Which languages should the TMS support?

   — Unicode (ISO 10646) is the recommended encoding standard for interoperability with other systems and to support the widest possible range of languages. All newly developed TMSs shall be Unicode-compliant.

   — The choice of UTF8 or higher depends on the languages to be implemented and the available operating system.

   — If the TMS includes bidirectional languages (such as Arabic or Hebrew), bidirectional input methods and display formats shall be supported.

6. How many languages should the TMS support, both in terms of content and in terms of the interface?

   — The design of a TMS for monolingual, bilingual, and multilingual terminological resources will be quite different.

   — Monolingual databases are developed primarily as reference tools for writers and other users of the language, whereas bilingual and multilingual databases are developed primarily for translators and multilingual communicators. Monolingual databases typically record more information about term usage and tend to place more importance on definitions. They also more frequently record information about related terms and synonyms. Bilingual and multilingual databases, on the other hand, typically record context sentences, a data category that is often omitted in monolingual resources. These databases should obviously also feature the full range of data categories for translation-oriented terminology: language identifiers, transfer comments, degree of equivalency indicators, regional identifiers, translation-specific comments, and so forth.

   — If the same TMS is intended to be used for monolingual, bilingual, and multilingual applications, then different views should be created for each specific user type and activity. Each view can hide unwanted fields and emphasize important ones.

   — For a bilingual or multilingual TMS, designers should also consider whether there will be one specific source language (pivot language) or whether database search functions will be bidirectional or multidirectional with respect to source and target languages. For example, when the source language is fixed, the user cannot set the search language, and the user interface can be designed such that the data on this source language always appear at the top of the screen. But in a bilingual TMS that supports search on either language, the user interface should allow the user to select the source language, and the search results could dynamically adjust to display the selected source language at the top. In a multilingual TMS that allows search on any language, the user interface should allow the user to select the source language plus one or more target languages. In any case, if it is determined that a majority of users will typically choose one specific source language, then that source language should be set as the default search language to maximize efficiency, but users should be able to change that language if desired.

   — Within a multilingual, concept-oriented TMS, designers should consider how to handle situations involving quasi-equivalent concepts: a term A can be considered equivalent to B in one language and C in another, but the concepts underlying B and C may not be fully equivalent. A decision should be made about how to handle such cases of incomplete concept equivalence, as this decision can impact the core design of the TMS. These partially equivalent concepts may either be included in the

Documents
Document Sharing Hub

© ISO 2012 – All rights reserved

Full document details: BS ISO 26162:2012