

Terminology Starter Guide

Contents

About.....	4
Who should read this document?.....	5
Introduction.....	6
Effective communications.....	6
What is a term?.....	6
What is terminology management?.....	7
Why manage terminology?.....	7
What is a terminology management system?.....	7
Settings where terminology is managed.....	9
Commercial production settings.....	9
Government.....	9
Content development.....	10
Localization.....	10
Benefits of terminology management.....	11
Reducing costs and time to market.....	11
Improving quality.....	11
Strengthening brands and protecting intellectual property.....	11
Preserving know-how.....	11
Common applications of terminology.....	12
Improving terminology in translations.....	12
Checking terminology in the source language.....	12
Modelling taxonomies and ontologies.....	12
Creating and managing a termbase.....	13
What is a termbase?.....	13
Concept-based approach.....	13
Data categories.....	14
Dealing with terms.....	14
What terms should be defined?.....	14
Guidelines for definitions.....	15
TermBase eXchange (TBX).....	15
Terminology management systems.....	16
Broad types.....	16
Key features.....	16
Single repository.....	16
Concept orientation.....	17

- Term autonomy..... 17
- Data elementarity..... 17
- Workflows..... 17
- Quality assurance..... 18
- Users..... 18
- Reporting..... 19
- System integration..... 19

- Terminology extraction.....20**
 - Statistical versus linguistic approaches..... 20
 - Observing terms in context..... 20
 - Other features.....20

- Setting up a TMS.....21**

- Common risk factors.....23**

- People.....24**
 - Overview of the roles of team members.....24
 - Executive sponsor..... 24
 - Steering committee..... 24
 - Project manager..... 25
 - Terminologist..... 25
 - Content developer..... 25
 - Editor..... 26
 - Translator..... 26

- Conclusion.....27**

About

Developed by: Terminology for Large Organizations (TerminOrgs)

Established in 2011, Terminology for Large Organizations (TerminOrgs) is a consortium of terminologists and other communications professionals who promote terminology management as an essential communications strategy in large organizations.

This group is a forum to discuss and develop guidelines and best practices for large-scale terminology management. Our mission is to raise awareness about the role of terminology for effective communications, knowledge transfer, education, risk mitigation, content management, translation and global marketing, with a focus on large organizations. TerminOrgs represents stakeholders of terminology standards and tools. We work to determine and promote the economic value of managing terminology.

To learn more about TerminOrgs and our activities, please visit www.terminorgs.net

Document editor: Kara Warburton

Version

December 2016. This document is updated periodically. Please visit <http://www.terminorgs.net> to download the latest version.

Copyright

© Copyright Terminology for Large Organizations (TerminOrgs) 2016. All rights reserved. This publication may not be distributed for commercial purposes, reproduced, modified or translated without the explicit prior written permission of the copyright owner. For permission, contact info@terminorgs.net. The TerminOrgs logo is the property of Terminology for Large Organizations (TerminOrgs). These materials are provided by Terminology for Large Organizations (TerminOrgs) for informational purposes only, without representation or warranty of any kind, and Terminology for Large Organizations (TerminOrgs) shall not be liable for errors or omissions with respect to the materials.

Who should read this document?

This guide provides an introduction to terminology management as it applies to large organizations such as global enterprises, governments, non-governmental public organizations, and large localization service providers. It is intended for people who have been requested to manage terminology in such environments. It is not a comprehensive manual about terminology management.

Introduction

Effective communications

Effective communications is a goal of all organizations that deal with the public, commercially or otherwise. This includes businesses, enterprises, public institutions, NGOs, governments, and any other type of organization. When the organization operates in different linguistic communities, requiring different languages, the goal of effective communication requires a proactive approach that includes terminology management. These types of organizations are characterized as "global."

Terminology is normally managed and published in a database, referred to as a "termbase", which is used by all employees and suppliers, including translation services, communications, public relations and marketing. Termbases are primarily used in connection with translation memories and machine translation tools, but they can also be combined with other knowledge management systems such as management information systems, document management systems or content management systems, controlled authoring software, and so forth.

What is a term?

There is no consensus about what constitutes a term, as opposed to an ordinary word or other type of expression. According to conventional theory, a term is an expression that designates a particular concept within a given subject field. In practice, terms do not always designate concepts in a given subject field. To support the communicative aims of large organizations, the notion of a "term" extends beyond the conventional view to include any expression that, if it is managed according to the methods outlined in this document, brings some benefit to the organization such as improved communication and reduced translation costs. This includes, sometimes, words from general language, marketing slogans, short sentence fragments, and so forth.

Nowadays, terms are typically recorded and managed in a terminology database (termbase).

Thus, terms in a termbase can include words and expressions that have a specialized meaning (technical, scientific, market-specific, political, etc.), or that are prominent in customer-facing materials (slogans, names of products, features, programs, parts, frequently-occurring words, labels on user interfaces, packaging, etc.).

The traditional view holds that terms convey special meanings as opposed to "words", which convey general meanings. However in commercial environments, any word or expression that is deemed to be important for the company's business and communications qualifies as a "term" subject to inclusion in the company termbase, regardless of whether its meaning is "special" or not.

Terms should be regarded as the linguistic expressions for individual concepts. A "term" in fact is a binary relationship between a linguistic form (one or more words) and the concept (the meaning). In most languages, a given linguistic form (word) can have more than one meaning. In this case, the word is actually considered two different words and they are referred to as "homographs".

For example, "copy" has at least three different meanings: (1) a duplicate of an item, (2) a piece of marketing text, and (3) the action of making a copy. Two different nouns and a verb. The word "port" has many meanings: a place for boats, a strong wine, the connection point on a computer, and the action of converting program code to operate on a different operating system. In other languages, the equivalent words are unlikely to be the same as they are in English, because the meaning is different.

In terminology management, homographs are considered to be, and treated as, different terms. Each will be recorded in its own entry in a termbase. Never would the four different meanings of "port" be documented in the same entry. This is what is meant by terms having a binary relationship between the form and the meaning. Each form-meaning pair is a distinct term. So in the example of "port", we are actually dealing with four terms, not one. This conceptual basis for managing terminology is fundamental.

What is terminology management?

Terminology management is the set of activities carried out to ensure that the correct terms are used consistently across the organization, in support of end-to-end product development, communication, translation/localization, and distribution. First, it involves collecting the terms of interest in the organization. Second, it encompasses documenting these terms with the appropriate information (metadata), such as definitions, subject field, product associations, usage guidelines, and part of speech. Third, it usually also requires a review of the terms to decide on preferences -- to help guide writers and translators to use language consistently. Fourth, it comprises the distribution and dissemination of this information as reference materials for writers and translators and as linguistic resources for enterprise systems including, but not limited to, content management and authoring tools, translation tools, product taxonomies, and search optimization.

Why manage terminology?

Managing terminology can positively impact profitability, customer relations and the overall image of an organization.

If terminology is not managed, your staff, business partners, and suppliers will decide for themselves what term to use to communicate about your products or services, and this could be contrary to the image you would like to portray as an organization. Problems that can and will occur include the use of synonyms (different terms used for the same concept), inconsistencies (in capitalization, hyphenation, spelling, and so forth), acronyms and abbreviations that are not explained, terms that have negative or inappropriate cultural connotations, marketing terms that do not convey the strongest positive image, terms that create problems for search and retrieval, terms that are unnecessarily difficult to translate, and even terms that convey the wrong concept altogether. These problems make your company look unprofessional and cause customer confusion and dissatisfaction. Some even lead to misuse of products which can cause accidents and harm the user. In addition, terminology problems within a company add costs in several ways.

At the research and development stage, the use of different terms for core features or functions can lead to misunderstanding among workers. Errors can occur, and some production tasks may even need to be repeated as a consequence, often at great cost.

After a product or service has been developed, the informational and marketing content is produced, then translated. There is often a disconnect between the marketing department and the product development department. Each has its own team of writers. Inconsistent and conflicting terminology between marketing and development content concerning the same topic (product or service) is a common problem. A centralized termbase is a tool that helps to ensure consistent and appropriate use of language throughout the organization. Without a termbase, language problems are left to editors to detect based on their own internal knowledge. Many inconsistencies and problems are undetected, and are then repeated unknowingly by translators in the translated versions. Furthermore, the editing stage is the very end of the content production cycle, after nearly all the content for a product has been produced. At this stage, the problems are multiplied many times over and the cost of fixing them is substantial.

Proactively establishing consistent terminology across the organization is not just an exercise in good quality writing and translation, it also improves the effectiveness and efficiency of a business as a whole.

What is a terminology management system?

A terminology management system (TMS) is computer software that helps you to store and retrieve terminological information.

It stores terms, associated and illustrative information such as definitions, contexts, and images, classificatory information such as subject fields and product associations, guidelines such as usage recommendations and indications of preferred synonyms, as well as administrative information such as creator, creation date, and change history. Most terminology management systems are customizable -- you can decide which fields you need and how to arrange them.

A TMS provides search functions and filters for the purpose of retrieving information. You can search for terms, but you should also be able to search for information in other fields in the database; the latter is known as a full-text search. Searches are carried out directly in the TMS, i.e. within its own user interface (browser or rich client application). Some systems are also integrated into other language technology applications, such as translation editors or authoring tools. These tools may then also allow you to execute searches in the TMS and display the results within their own interface.

Settings where terminology is managed

Terminology is managed in various settings; several settings where it is managed in a large scale are presented in the next sections.

Commercial production settings

In settings where productivity in content production and translation is the driving goal, such as a global enterprise or other large organization, terminology management needs to be a proactive, multi-disciplinary and integrated process. Proactive terminology management aspires to intercept terminology at the time of inception, when concepts are created and named. Various people in the organization participate in terminology discussions, decisions and implementations from R&D to marketing, from developers to product managers, from writers (content developers) and editors to translators, and from sales force to support teams. Terminology management is an integral part of the business' workflows and is built into formal objectives and performance measurements of the key stakeholders and users, such as content developers and translators.

Ideally terminology is established early in the development of a product or service, and put into the termbase. Content producers, editors, and marketing writers use the termbase as a resource for checking terminology. Standardized target-language equivalents are added by the target-language terminologists and other translators. kWith this distributed approach for building a centralized language resource, key people from across the organization working together to prescribe consistent corporate language. The aim is to form an identity, to establish clear communication, and to thereby realize savings in time, resources, and money.

Rather than limiting terminology management to an isolated activity in the translation department, the successful organization incorporates it as a holistic approach that unifies corporate activities such as design, quality control, branding and marketing, search, text mining and analysis, content management and controlled authoring, as well as human and machine translation.

Government

Good terminology management in government has some of the same objectives as in product localization or commercial communications: reducing translation time and assisting translators in producing consistent, high quality translations. However, even before the translation step, good terminology management also helps public servants write better, clearer documents in the first place.

Wasted time is wasted taxpayer money and therefore a government should be clear and coherent when communicating internally and with its citizens. For governments, particularly those with more than one official language, consistent terminology is vital - when writing laws, for example. Imagine what could happen if the terminology used in a country's constitution was inconsistent within one language and then translated inconsistently into others, with each version having the same legal weight. One arm of a government cannot accurately and effectively communicate with another without consistent terminology. Governments cannot effectively deliver services to their "clients" (the citizenry) without consistent terminology. Multiply the number of misunderstandings (including potentially grave ones) that can be made about one concept in one language by the number of official languages and it is clear that terminology management is important in government. Over time, poor translations could give the impression of carelessness, and even result in alienation of voters who speak a particular official language. Managing terminology well is in the best interests of governments everywhere, easing internal communication and also helping to present a unified message to other countries and international organizations. Governments that have realized this, such as Switzerland, France and Canada, have dedicated resources to terminology management.

In countries that have more than one official language, it is a citizen's right to receive information in the official language of his or her choice. Terminology management is therefore a key component in a government's legal responsibility to provide linguistic services. Terminology management also serves a role in support and preservation of endangered languages.

Content development

In many corporate environments that support terminology management, it is the technical writer who is on the forefront of terminology development. In this environment, the technical writer, the product manager, and the terminologist identify new terminology that emerges as products are being developed: product user interfaces, error and system messages, internal product design specifications, vocabulary to be shared in everyday communication, and customer documentation. The terminologist adds the new terms to the company termbase. Of ultimate importance is agreement on terms and definitions for consistent use in all many corporate environments, it is the content developer who is at the forefront of terminology use. In this environment, the content developer, the product manager, and the terminologist identify new terminology that emerges as products are being developed: product user interfaces, error and system messages, internal product design specifications, vocabulary to be shared in everyday communication, and customer documentation. The terminologist adds the new terms to the company termbase. It is then necessary to reach agreement on the terms that are allowed in all corporate collateral. The obvious beneficiary of such due diligence is the customer.

The content developer must be aware of emergent terminology during the early phase of a product's development cycle. A clear identification and understanding of new terms is critical to the writer who develops detailed conceptual topics in the documentation.

The writer is often an active member of a product development team, working with R&D to develop clear and consistent terminology within and across product teams in the corporation. The writer must be sufficiently skilled to recognize the difference between everyday language words that do not warrant any special treatment or controls, and terms and other expressions that, by virtue of some property such as visibility, difficulty, or frequency, need to be researched, documented, and shared with content producers and translators.

In some environments, the writer participates in the process of entering and managing product-specific terminological data in the organization's termbase using a particular TMS. The writer's level of access to particular sections in the termbase can be appropriately controlled. For example, the writer might have the ability to add terms and definitions to a particular section of the termbase, where the data added is automatically given a status of "unprocessed." Later, a terminologist or an editor checks the information and changes the status to "finalized" (approved). This is just one common workflow in a TMS.

Ideally, the writer provides terms and descriptive information about terms in the source language, and localization specialists (translators) then add corresponding terms and information in the target languages. The result of this collaboration is a multilingual termbase.

It should be possible to export client-ready glossaries from the TMS by using an export function. These glossaries can be included in the product documentation that is provided to customers.

All the organization's employees, and even external partners, should be able to access the terminology information in the termbase. For this purpose, the terminologist or system administrator should provide tailored views. The TMS should also be sophisticated enough to filter the information according to user type, for instance, to hide confidential or potentially sensitive information from external users.

Localization

The term "localization" is now commonplace to refer to translation work in computing environments. Good terminology management reduces translation time and assists translators in producing good quality and consistent translations. All this helps to significantly reduce localization costs; translators spend less time researching or querying terminology and quality issues are minimized. Consistency in source terminology enables previously localized material to be reused to a greater extent; leverage of translation memories is higher when the source text is consistent. Some concepts can be represented by several different terms (synonyms). Bear in mind that most localization projects involve translation to multiple languages – so standardizing terminology in both the source and target languages facilitates accuracy and establishes a consistent approach. Reducing the need for rework by getting it right first time, i.e. by providing a termbase as a reference to content producers and translators, not only improves communications and establishes a good brand image, but it also saves money.

Benefits of terminology management

Managing terminology reduces production costs (in all languages), improves customer relations (in all markets) and protects and strengthens the overall image of an organization.

Reducing costs and time to market

Well managed terminology increases the ability to reuse and retrieve information. Editors, translators and other specialists spend less time searching for correct terms, definitions and equivalents, hence increasing their efficiency. Furthermore, the consistent use of terminology throughout the document life cycle leads to a reduction of errors and inconsistencies in the source language and their propagation in translated versions, which in turn reduces the cost of revising and reprinting documents and helps shorten time to market. After-sales costs can be reduced if misunderstandings or errors are avoided by producing comprehensible documents.

Improving quality

Managed terminology improves the quality of products and documents. This in turn increases the usability of products and thereby customer satisfaction which could lead to increased market share. Information is easier to understand if the terms are used consistently in all languages. Documents that are intended for clients will contain terms suitable for their target group, specialized expressions will be clearly defined, confusing jargon will be avoided, and abbreviations will be appropriately explained.

Strengthening brands and protecting intellectual property

Successful brands create a positive impression, a positive customer experience, which in turn leads to customer loyalty. Key marketing terms need to be strong, clear, and consistent. In the era of global one-voice strategies and virtualization of client contacts, managing the terminology for all markets and languages is essential.

Preserving know-how

Concepts are the basic units of knowledge. Terms are the linguistic representation of concepts, whereas definitions can be seen as explanations of concepts. These two elements, terms and definitions, are key components in termbases and therefore tremendously precious to a company. With every specialist who leaves the company, a portion of knowledge is lost. Some of this knowledge can be retained if it is stored in a termbase. Human Resources and senior employees can use termbases to explain key concepts to new employees, thus making sure that continuity of knowledge is preserved.

Common applications of terminology

Improving terminology in translations

Translators benefit from being provided with some assistance with respect to approved or recommended terms and other expressions. Therefore, the translation tool used should be able to send the sentence currently being translated to the TMS so that known terms are automatically identified and highlighted, and the translators can then choose the approved target terms.

Furthermore, a translation tool may access a termbase and compare the terms actually used in the translated text with the ones that are stored in the termbase. If a translator has used terminology other than that which is prescribed in the database, the system raises an alert that can be investigated.

Checking terminology in the source language

Authoring tools can look up terms in the termbase to verify whether the right terms are used. If the termbase contains not only the “approved” terms but has also the undesired “deprecated” terms classified as such, the authoring tool can evaluate this information, allowing it to warn the writer about terms that should not be used.

Modelling taxonomies and ontologies

Many organizations develop taxonomies or ontologies to structure their knowledge, product lines, and so forth. While a taxonomy or ontology models and stores the relationship between concepts, terms form the navigable “handles” for browsing the data. Therefore, taxonomies and ontologies store terms in a very structured manner. The terms that are used as labels in taxonomies and ontologies should be standardized within the organization. There is therefore a strong relationship between ontologies, taxonomies and termbases, and their development should be synchronized.

Creating and managing a termbase

What is a termbase?

According to ISO 30042:2008, a termbase is “a database comprising information about special language concepts and terms designated to represent these concepts, along with associated conceptual, term-related, and administrative information.” In commercial settings, however, it should be noted that a termbase is not restricted to “special language concepts”; it may contain any lexical unit that needs to be “managed” in order to produce quality, consistent content.

Termbase management is a combination of terminology work and database administrative tasks supporting the systematic collection, description, processing, presentation, and distribution of information about the terms and other various linguistic units used in the organization.

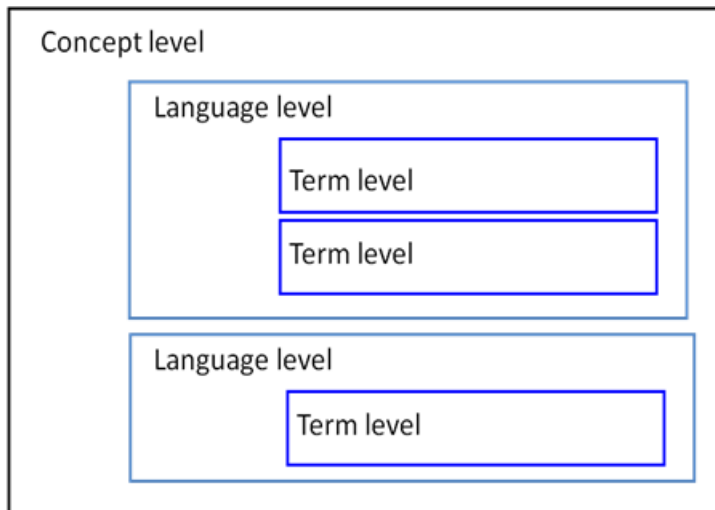
Concept-based approach

A concept-based approach to terminology management requires all terminological information pertaining to one concept to be handled as a single terminological entry. In a concept-based system, data is organized around the meaning rather than the linguistic form (the term). All terms, variants, and translations that designate a single concept, as well as all descriptive and administrative data pertaining to that same concept, are stored in a single terminological entry.

This approach is what distinguishes terminology and terminology management from lexicology and lexicography, which is the practice of developing dictionaries (terminology management is sometimes called “terminography” by extension with the term “lexicography”). Dictionaries are “word-based” in contrast to termbases. Each entry in a dictionary is centered around a word, and describes all the various meanings of that word. Except in dictionaries that are thesaurus-like, the entries do not normally include synonyms; one has to look elsewhere in the dictionary for those.

In a termbase, all synonyms must be included in the same entry. If a term has more than one meaning, then a separate entry is required for each meaning. This is why in the navigation list of its terms, there are cases where the same term is listed more than once in a termbase. They may look like duplicate entries, but unless the meaning of the entries is the same, they are not. If the meaning is indeed the same, then the entries are true duplicates and they should be merged into one entry. These types of duplicates are called “doublettes”. Doublettes often occur after a new set of terms is “imported” into the termbase. Eliminating doublettes is one of the main maintenance tasks of the terminologist.

Within the concept entry, there are three hierarchical sections: concept level, language level, and term level.



Data categories

Data categories are the different types of information recorded in termbases, in the different fields of the entries, such as definitions, parts of speech, and usage notes. They guide the organization and arrangement of information in each entry. The data categories determine what types of information go into each entry, what information is mandatory vs. optional, the format of the information (i.e. free text, predefined pick lists, or Boolean values), and the input method (e.g. manual vs. system generated).

An inventory of data categories for recording terminology has been defined by ISO Technical Committee 37. The Data Category Registry, ISocat (<http://www.isocat.org/>), is a free online resource for specification and management of data categories and data category selections for language resources. The data categories in TBX-Basic are the most popular ones used in the localization industry, as determined by studies conducted by the Localization Industry Standards Association in 2001 and 2005.

Dealing with terms

What terms should be defined?

Preparing definitions is time-consuming, and it is not practical to write a definition for each term in your termbase. Most terminology processes in production-centred environments do not emphasize the inclusion of definitions in the termbase. Therefore, you need to decide which terms need definitions, and which terms are self-explanatory. It is generally agreed that you should define terms that are not found in general dictionaries, domain-specific terms (terms that have a special meaning in a field), terms that have a unique meaning or usage in your organization, and any terms that have caused confusion in the past. Basically, provide a definition only if you think someone will need it.

For instance, if a term has multiple meanings in the organization, then each meaning should be clearly defined since each distinct usage of the term is in fact a unique concept and is likely to have its own translation. If a term is similar in meaning to another term, the two should be clearly defined so that the nuance between the two can be properly reflected in the translations. Terms that have multiple meanings (polysemes) should be defined.

If the termbase is used to generate glossaries for clients, consumers, or for the public, then you need to define terms that may be unfamiliar to them. For a product glossary, focus on the frequently-occurring product-specific terms such as the names of product features and underlying technologies. Ensure that all acronyms are expanded to their full forms and defined.

The choice of which terms to define also depends on your goal. For example, the terms you would choose to define if your goal is to make your company communicate better internally will differ from those you would choose to define for research in a particular field, for the engineering and development of a product, or for educational purposes.

As an alternative to a definition, you can include what is called a “context” -- this is a sentence that contains the term.

Terminology databases without definitions or contexts are not much better than simple lists of terms, the value and reusability of which diminishes quickly over time.

Guidelines for definitions

A definition should answer the question “what is it?” A good definition specifies the essential and delimiting features of the concept, so that you can tell what it is, as well as how it differs from other related concepts. The definition should clearly delineate one concept from another.

When writing a definition, start it with an anchor word that refers to the class the concept belongs to, usually a generic or superordinate word of the same part of speech (noun, verb, adjective, etc) as the term. (Ex: skirt: An article of clothing that...). Then, add delimiting characteristics (Ex: An article of clothing that hangs from the waist or hips and covers the body below the torso.)

A good definition is a succinct, clear statement, usually in the form of one sentence, that describes the meaning of the term. It should not contain additional information, explanations, or notes. These types of information should be included in the entry in their own dedicated fields. Write concise definitions that include the essential characteristics of a concept in a definition, and record other useful information about the term in other fields so it is not lost.

TermBase eXchange (TBX)

Term Base eXchange (TBX) is the open, XML-based standard for exchanging structured terminological data. TBX has been approved as an international standard by ISO (ISO 30042) and is available from www.tbxinfo.net.

A TBX variant called TBX-Basic was developed by TerminOrgs and is available from www.terminorgs.net. TBX-Basic is a lighter version of TBX. It is well suited for any language application that requires an approach to terminology management that is based on core principles and common requirements.

Terminology management systems

Successful terminology management requires the use of a terminology management system (TMS) that can adapt to constant product development, tools and infrastructure improvements, and organizational changes. The key principles that support these requirements and that should guide the selection of a TMS are summarized in the following sections.

Broad types

When acquiring a TMS, there are three categories of solutions to consider.

Existing technologies and software

Spreadsheets and word processing applications are often used to collect lists of important terms. They are frequently used prior to the adoption of a TMS.

Advantage: Easily available.

Disadvantage: Not able to do proper, concept-based terminology management; not integrated into translation tools; does not easily accommodate concurrent user or access rights.

Accompanying module in a larger software solution

Translation tools, authoring tools, enterprise search, and taxonomy management tools have a requirement to store monolingual or multilingual terminology. Therefore they often include some terminology management functions. The available features and level of sophistication range from very simple (flat term list) to very sophisticated (feature-rich, concept-oriented).

Advantage: Integrated into existing software in use.

Disadvantage: Often not able to provide proper terminology management, as features are driven by the parent software (the terminology management functions in translation tools or authoring tools often have a smaller scope than a proper TMS).

Dedicated terminology tool

Full-featured software with the primary purpose of storing, maintaining, and distributing terminology data. It may be server and/or web based, or alternatively a stand-alone software.

Advantage: Very appropriate for terminology needs.

Disadvantage: Integration with translation or authoring solutions may require a separate “integrator” software, or an export/import process.

Key features

The following sections describe the key features to look for in a terminology management system that make it suitable for managing terminology in a large production environment.

Single repository

All data is managed in a single database or repository. Different types of terminological units (such as single and multi-word terms, acronyms and symbols) are identified using data categories, and users can access the sections of the terminological database that are relevant to them. If a single repository is not possible, then the TMS should allow for some kind of simultaneous search and access to all termbases.

Concept orientation

The TMS should be concept-based, as described in *Concept-based approach* on page 13

Term autonomy

It should be possible to describe all terms with the same degree of detail (that is, using all the same fields in the system). All terms are recorded in autonomous and repeatable blocks of data categories within a terminological entry.

An example of the violation of this principle is when the termbase contains a field called **Acronym**: where one inserts the acronym of a term, i.e. putting terms and acronyms in separate fields. To respect term autonomy, both the full form and the acronym should be put into their own **Term**: field and the property of the term being an acronym or a full form should be indicated in a **Term type**: field.

Data elementarity

All information is separated into dedicated fields so that there is only one type of information per field. For example, it would be incorrect to include explanatory information in parenthesis after the term in the term field, or to provide both context and definition in the definition field.

Workflows

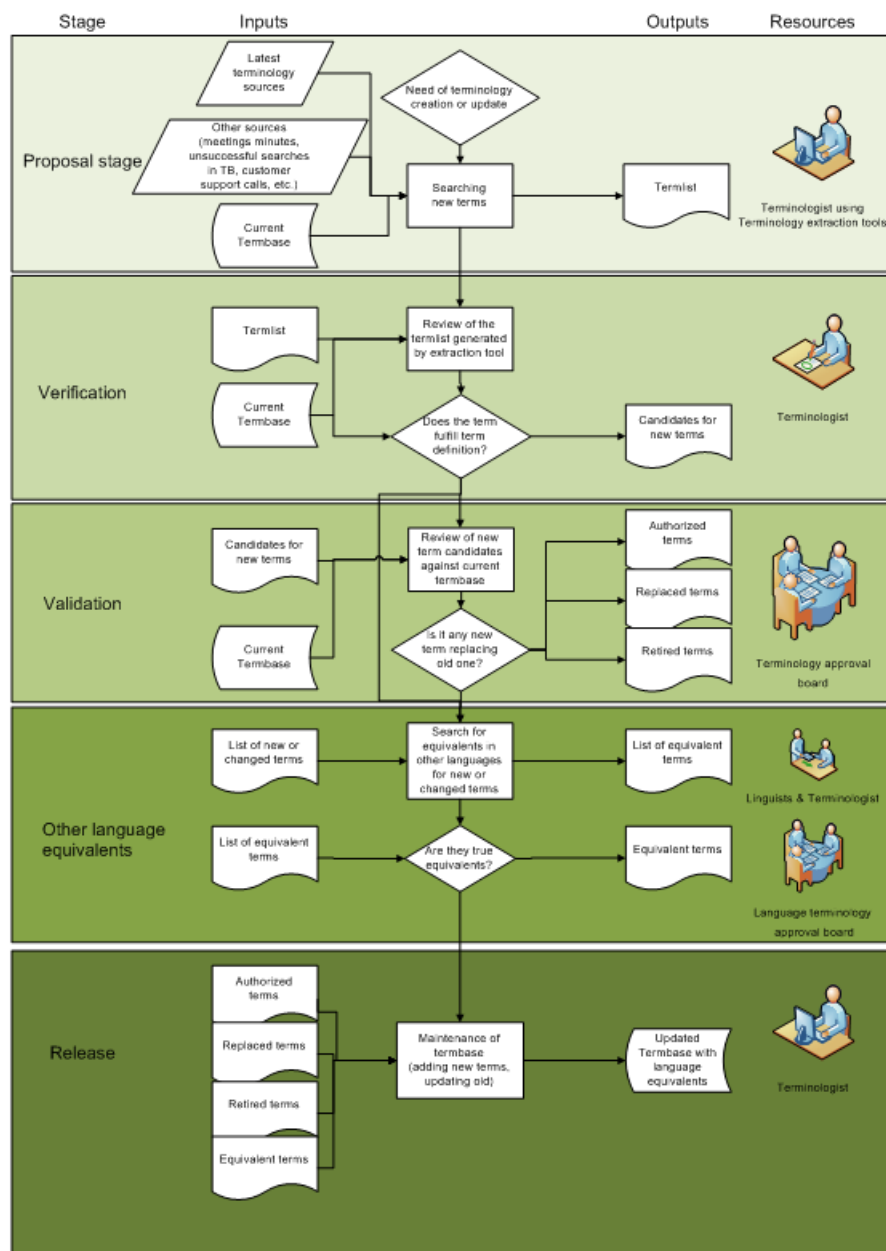
The process of creating a new term entry commonly includes these steps:

- Proposal
- Verification
- Validation
- Release
- Addition of equivalents in other languages

A term life cycle includes the creation and validation of new entries, evaluation of synonym sets and assignment of status values to rank terms according to style and usage objectives, and maintenance tasks such as adding complementary information and merging doublettes. Terms that are incorrect, invalid or obsolete are not removed from the termbase; rather, they are identified as such by using a data category that indicates their status as deprecated or no longer used.

The terminology management system should provide workflow-type functions that support this process.

The following illustration outlines the common steps in a terminology workflow:



Quality assurance

Quality assurance is needed on three levels:

- Verifying that new terms (neologisms) follow any existing brand strategy, copyright regulations, usability guidelines, and sound term formation.
- Verifying that information in each terminological entry is entered accurately, correctly and according to adopted standards.
- Governance mechanisms are required to establish preferred terms and provide other usage guidelines.

Users

To support both the usability and quality of a termbase, the TMS must support access by multiple users, as well as safeguard the information that it contains.

- The termbase should be available to everyone in the organization and allow concurrent users. Different users may have different types of access (read-only, read-write) based on their roles and needs.
- Multiple concurrent users demands a security model that distinguishes and restricts access rights based on a combination of criteria:
 - Data level: users are granted privileges to either the whole entry or only a portion of the entry
 - Language: users are granted privileges on a language basis
 - Roles: users are granted privileges according to their role (e.g. casual user, term submitter, term reviewer, term approver)
 - Subject area: users are granted privileges according to their areas of expertise
 - Organizational structure: users are granted privileges according to their position in the organizational structure (e.g. internal vs. external users, members of particular departments, etc.)

Reporting

The three common types of reporting needs for terminology management are described below:

- Status reports typically cover quantity of terms and entries, including information on number of new, changed, and deprecated entries (globally, per language, or per project, etc.).
- Maintenance reports are used to detect quality issues. Quality issues may include invalid data category combinations (for example a definition written for a noun, but part of speech set to ‘verb’), empty fields, possible doublettes, and spelling and grammar errors (in text fields).
- Terminology change reports provide information on terminology changes and updates. These can be used to check other resources (e.g. documentation, translation memories) that may be affected by terminology changes. Change reports are also useful for informing affected teams in the organization of any changes to key terms.

System integration

In most large organizations, terminology in a termbase needs to be integrated into other enterprise systems, such as authoring and translation tools. The TMS should have a documented API and it should support file formats that are compatible with other tools in areas such as:

- Product development
- Document authoring and knowledge management
- Quality assurance
- Translation
- Term extraction or harvesting
- Customer management and care

Terminology extraction

To help collect terms for the termbase, using an terminology extraction process can be extremely useful. It is often too time-consuming to manually identify the important terms in an information set. Running a terminology extraction tool on a text or group of texts (corpus) automatically produces a list of words and word combinations that might be terms, called “term candidates.” A person (often a terminologist) will then go through the list to determine which of the words and phrases would be useful to record in the termbase. The terms selected by the terminologist will often become the focus of further terminological research, including translation into the required languages.

With a term extraction tool, you can find all occurrences of a term and information about a term much more quickly and systematically than you would by searching manually.

Some key features to look for in term extraction tools are described in the following sections.

Statistical versus linguistic approaches

Some term extraction tools find the terms by using statistics only (which words or word combinations occur most often). Because the approach is purely statistical, these tools often work equally well on any language.

Other tools integrate linguistic rules to help determine which words or word combinations are likely to be terms -- usually noun combinations, as opposed to function words such as prepositions and articles, phrases or random strings of words. Because these tools rely on linguistic rules, they tend to only work on one language or a limited number of languages. However, the results are generally better than the purely statistical tools.

Another type of tool can process bitexts (the pair of a source text and its translation, aligned), and extract terms in both languages: source language terms and one or more likely translations for each. Some of these tools require a person to validate the suggested translation of a source term, while others can be set to select what is statistically the “most likely” match.

Observing terms in context

Most terminology extraction tools allow you to view all or several occurrences of the term in context. Being able to browse some sentences in which that term occurs makes it easier to first verify that the term exists and, in the case of multi-word terms, is properly delimited, and also to locate other information, such as definitions, synonyms, and examples of usage.

Other features

Other important features to look for in a term extraction tool include the following:

- Extract both single-word terms and multi-word terms
- Indicate the frequency of occurrence of the term
- Indicate the part of speech of the term
- Indicate the name of the file from which each context sentence was extracted
- Allow an exclusion list -- to exclude unwanted terms from the output

Setting up a TMS

The following are the key steps for setting up a terminology management system (TMS)

1. Identify what organizational goals can be served by a terminology management process and a termbase: Increased efficiency and productivity ()? Improving communication internally/externally? Cost savings? Begin to develop a business case around these themes.
2. Identify what resources you already have: Persons who might be able to assist in developing and managing the termbase, available computer hardware, servers, networks, etc.
3. Get the terminology expertise you need: Hire or train someone who has a suitable background to be the main resource for terminology for your corporation. Good candidates include persons with a formal education in languages, linguistics, content development (technical or marketing writing), and translation.
4. Identify the information that you require for each term (data categories, for the fields in the termbase). A good foundation would be the data categories in TBX-Basic.
5. Decide which networking infrastructure you need. Online web-based terminology database, or a client-server application? Central control, or open like a wiki?
6. Consider the terminology software available and find the one that meets your needs. Some questions you might ask include:
 - How many people might need to connect simultaneously?
 - Is there a limit to the number of records can it handle?
 - What views does it offer (dictionary style vs. single-concept)? Can you create custom views?
 - Are there automatable workflows? How good are the search filters?
 - Can you control read/write privileges?
 - Is the cost within your budget? Enquire about both initial costs and ongoing maintenance and upgrades of the software.
 - If you deal with many languages, does the tool treat all languages equally? Is there support for language variants, if you need them (such as Canadian French and European French)?
 - Which technology suppliers are you already dealing with? Can they offer any solutions? What do they have?
 - Are import and export formats compatible with the various tools the TMS will need to integrate with? Does it support TBX? (Note: TBX support should be independently tested and verified; marketing claims of TBX support are not always reliable.)
7. Create an implementation plan. Include a training plan for users, and a communication plan to promote the new process and tools.
8. Define a maintenance plan to review and update the content of the database as new terms and other information are added. Remind users on a regular basis that the terminology service is there to help them and that they are encouraged to use it.

An important aspect of the successful implementation of a termbase is to engage as many people in the organization as possible. Ensure that you:

- Build awareness across the organization. In organizations where terminology management is relatively unknown, it is recommended that this phase of the implementation starts early in order to raise awareness and develop the necessary technical skills.
- Identify key stakeholders and recruit champions (stakeholder representatives who are willing to support the endeavour).
- Develop training modules that are customized to the different user types and scenarios.
- Reward active participation: For example, on a quarterly basis, select a terminology help request and calculate how much time and/or money one terminology search has saved the company (include multipliers such as number of language versions and number of text formats). Recognize the employee's efforts and savings to the company. Reward them if possible, and make that effort known to all employees. Show your employees how such a seemingly small contribution (like asking which synonym is to be preferred and used) saved the company money, and that their contribution is valued.

To generate the required interest and participation, ensure that terminology management tools and processes are rolled-out to all users and stakeholders. Make announcements. Create publicity around the TMS in internal communication channels and events.

To ensure widespread adoption and use of your termbase and TMS, ensure that the users have a voice with regards to changes and improvements:

- Collect user feedback and requirements.
- Identify improvements to processes.
- Develop new functionalities or customizations.

Common risk factors

There are a number of risk factors that need to be accounted for during the early stages of implementing a TMS. These include, but are not necessarily limited to the following.

Lack of sponsorship from executives

Management buy-in and understanding of the importance of terminology management to the company's overall communication strategy is critical. Without a committed sponsorship from key executives, it is nearly impossible to implement a terminology management process at all, much less to realize its full potential.

Lack of buy-in from stakeholders

Terminology management is a collaborative and collective effort. In addition to top-down backing from upper management, bottom-up support from people on the front lines is crucial. Terminology management is truly successful only when each person in the organization consults it when writing, translating, or at other times. A risk factor is tied to executive sponsorship in that a continuous support and participation from stakeholders is only possible when such participation and support is made part of their formal objectives and performance measurements, or when the behavior becomes ingrained in the corporate culture.

Disconnect from product development

Terminology management needs to be an integral part of the product development cycle, from inception to sale, not as an afterthought in the documentation or translation departments.

Competing projects

Identifying other projects in the organization that may compete for resources (such as budget or personnel) at an early stage provides an opportunity to build on or collaborate with existing efforts, and to avoid duplicate work in terms of time, tools investments, and resources.

Insufficient resources

Without sufficient resources for staffing and tools, a terminology project can stagnate and fail to produce the desired return on investment. A solid business case is often required in order to acquire the sufficient resources.

People

Overview of the roles of team members

Overall, having the right people involved in the design, development, and implementation of a terminology management programme and its associated TMS makes the difference in how well terminology is accepted and used in the organization. The right mix of software can be in place, but a successful deployment depends on leveraging the talent and expertise of the core team as well as the extended organization.

Terminology management entails considerably more than creating a termbase by using a TMS. It's the "people component" that drives the goals and processes. Furthermore, the identification of the right people who have the right skills who must be in place at the right time is required in order to achieve success. This section presents recommendations for staffing a terminology team.

The core members of a terminology team are the following:

- Executive sponsor
- Steering committee
- Project manager
- Terminologists
- Content developers and editors
- Linguist
- Globalization and localization expert

In many situations, a dedicated, full-time Terminology Management group is not always available to do the work. Instead, people are "on loan" from other groups to assume various responsibilities.

The wants and needs of various types of end users are gathered in advance of the TMS design process and at various points during the development and testing phases.

Executive sponsor

An executive sponsor is required to ensure that the project has high-level support. The executive sponsor is attuned to the goals and dynamics of the corporation. The executive has relationships with the key stakeholders who are involved in setting strategic goals. Also, the executive understands the challenges, constraints, and the economic environment of the corporation.

With this corporate knowledge, the executive knows how to manage such an innovative initiative at the corporate level. Using knowledge of other corporate plans and priorities, the executive can advise the team on timelines and budgetary limits. An executive sponsor can "open doors" at various corporate levels in order to promote acceptance of the project.

Steering committee

The steering committee is a cross-disciplinary group that sets the vision and strategy for implementing a TMS within the corporation. For example, in a software company, the steering committee might consist of one relatively senior-level person from each of these divisions:

- Globalization and Localization
- Publications
- Research and Development
- Marketing
- Quality Assurance

- Technical Support
- Information Management Systems

Representatives from the key product lines may also be required. This team meets regularly, usually monthly or bi-monthly, to drive the terminology deployment project to completion.

Project manager

A good project manager is the key to the success of the TMS. Ideally, the project manager understands both the corporate climate and the technology. The project manager is also committed to success, has exceptional communication skills, and has management experience. The project manager is ultimately responsible for synthesizing and directing all aspects of the terminology management initiative. The project manager should have a reporting relationship to the executive sponsor in order to ensure direct communication.

Terminologist

Terminologists are personnel whose primary responsibility is to develop and maintain the termbase for two primary target audiences:

- External users of terminology (people who use the products or services). They need information about any potentially unfamiliar terms or uses of terms in the products and services.
- Internal users of terminology (employees): They need to clearly understand all the terms used in the organization, and use them correctly in their own communications.

Terminologists look for and record new terms (terms that have not yet been included in the termbase). They also manage existing terminology by updating records to include additional information or to improve existing information in the termbase as required.

They communicate with subject-matter experts to determine the appropriate terms for concepts and to develop accurate, precise concepts that are clear for global audiences; they populate and maintain the termbase with the appropriate terminological data; and they provide terminology advice and research.

Terminologists promote the importance of terminology management via timely and relevant articles about terminology topics on the corporate intranet.

While many terminologists work as described above, i.e., as language professionals specialised in the scientific study of the concepts and terms, there are also numerous examples of terminologists as generalists in the sense that their role is one of facilitation. They facilitate communication by providing tools and resources that allow people to communicate and speak with one voice. They liaise with stakeholders to develop processes, tools and resources that enable effective use of terms across the organization. They help design and implement terminology management systems and they provide relevant promotion and training.

Content developer

Terms in an organization are first used when content (documents, web sites, packaging, etc.,) is written in the source language by a content developer. After the content is written, it may be translated into various languages so that the product or service can be marketed world-wide. The correct use of terms in the source language is therefore critical because any inconsistencies or errors will be repeated in the translations, unknowingly by the translator. This is why managing terminology in the source language is extremely important.

It is the content developer's responsibility to use terminology that has been approved by the company. Content developers are therefore required to consult the termbase frequently. If the company uses a Controlled Authoring software, then this software should contain the approved terminology, and content developers can use it to verify that they are using the correct terms, just like a spell checker checks for correct spelling.

If the product or service includes any terms that may be unfamiliar to the user, the documentation provided should contain a glossary to explain them. The content developer should prepare this glossary by using the information in

the termbase. Any new terms should be added to the termbase. The terminologist, an editor, and possibly a technical reviewer (product engineer, subject-matter expert) should review the glossary terms and definitions. Any feedback should be incorporated into the termbase, and if possible, the terms and definitions should be exported directly from the termbase into the glossary format.

Editor

Editors work with content developers and terminologists to ensure that terms and definitions in source language in the termbase comply with grammar and style guidelines. As part of editing the content written by content developers, they also ensure that correct terminology is used. Editors should also check glossaries that are provided with products or services. They often have a broader view of an organization's overall content and content strategy than content developers, which allows them to identify terminology problems, conflicts, or challenges across project or product boundaries. This broader view makes them suitable to help the terminologist to review and approve source language entries in the termbase.

Translator

Translators add target-language terms to the termbase, after the content developer, editor, or terminologist has added the source-language terms. Experienced translators who have a good knowledge of the company's subject-matter can assume the role of the "target-language terminologist", reviewing and approving terms in the target language in the termbase, whereas translators with less experience may be restricted to "proposing" terms in the termbase, subject to final approval. A field in the entry records the status values necessary to handle this type of workflow (proposed, reviewed, approved, etc.).

Conclusion

Managing terminology is an ongoing activity. It is not a “project” that has an end. Terminology management needs to be part of the company strategy. Language changes, and these changes need to be managed – accepted and documented. Employees need to be informed about what the company requires and provided with adequate tools to enable them to comply with those requirements. Terminology management adds value to all communication within an organization and supports all levels of product and services life cycle and any accompanying communication.

Index

A

about this guide [4](#)
 applications of terminology [12](#), [12](#)

B

benefits
 branding [11](#)
 intellectual property [11](#)
 knowledge preservation [11](#)
 quality [11](#)
 reduced cost [11](#)
 reduced time to market [11](#)
 benefits of terminology management [11](#)

C

concept-based approach [13](#)
 concept orientation [17](#)
 conclusion, why [7](#)
 content developer [25](#)
 content development [10](#)
 contributors [4](#)
 corporate language [6](#)

D

data categories [14](#)
 data elementarity [17](#)
 definitions [14](#), [15](#)
 doublettes [13](#)

E

executive sponsor [24](#)

G

government [9](#)

L

lexicography [13](#)
 lexicology [13](#)
 localization [10](#)

O

ontologies [12](#)

P

people [24](#)
 people, team members [24](#)
 process [17](#)

production environments [9](#)
 project manager [25](#)

Q

quality assurance [18](#)

R

reporting [19](#)
 risk factors [23](#)

S

settings where terminology is managed [9](#)
 single repository [16](#)
 source language [12](#)
 steering committee [24](#)
 synonyms [13](#)
 system integration [19](#)

T

target audience [5](#)
 taxonomies [12](#)
 TBX [15](#)
 TBX-Basic [15](#)
 technical editor [26](#)
 term [6](#)
 term autonomy [17](#)
 termbase [13](#)
 terminography [13](#)
 terminologist [25](#)
 terminology extraction
 context [20](#)
 languages [20](#)
 linguistic [20](#)
 statistics [20](#)
 terminology in authoring [12](#)
 terminology in translations [12](#)
 terminology management [7](#)
 terminology management system [16](#)
 terminology management systems [16](#)
 terminology management system [7](#)
 TMS [16](#), [21](#)
 translator [26](#)

U

users [18](#)

W

workflow [17](#)