LISE

ICTPSP

LISE

ICTPSP

Project acronym: LISE

Grant Agreement Number: 270917

Project title: Legal Languages Interoperability Services

DELIVERABLE

D3.2 Report Workflow Adaptation for LISE

Dissemination level: PU

Version No. 6

31/07/2012

Project co-funded by the European Commission within the ICT Policy Support Programme.

Document information

Deliverable number:	D3.2
Deliverable title:	Report Workflow Adaptation for LISE
Due date of deliverable according to DOW:	M18
Actual submission date of deliverable:	31.07.2012
Main authors	Elena Chiocchetti, Natascia Ralli
Participants	
Reviewers	See revision history below
Work package	3
Work package leader	EURAC
Dissemination level	PU
Version	6

Revision history

Revision	Date	Author	Organisation	Description
1	29.06.2012	Natascia Ralli	EURAC	Initial draft
2	17.07.2012	Natascia Ralli	EURAC	Updated draft
3	24.07.2012 25.07.2012	Michael Wetzel Tanja Wissik, Barbara Heinisch-Obermoser, Vesna Lusicky	ESTeam UniVie	Section 7 and 8
4	25.07.2012	Natascia Ralli Isabella Stanizzi	EURAC	Final draft
5	30.07.2012	Barbara Heinisch-Obermoser, Vesna Lusicky	UniVie	Language revision
6	31.07.2012	Elena Chiocchetti, Natascia Ralli	EURAC	Final revision

Table of Contents

1. Objective of WP 3

WP3 “Legal Terminology & Workflow” focuses on the procedural content core of the LISE project: analysing, improving and streamlining existing workflows of terminology work in legal and administrative procedures, specifying terminology management procedures using the LISE tool suite on the service platform and finally, creating a description of best practices in the form of a guide for legal/administrative terminology work.

The present deliverable refers to **Task 3.2 - Report Workflow Adaptation for LISE,** as it explains how to adapt the terminology workflows described in deliverable **D3.1 - Analysis of existing workflows of terminology work** to the service requirements for the LISE online service on the basis of the tool suite made available to users.

NB. The **bibliographic references** have not been added within the text in order to make it more easily readable. Please refer to the bibliography listed in deliverable 3.1 for detailed references.

2. Background

2.1 Functions and types of terminology work

There are two main functions of terminology work:

a) Knowledge representation

Terminology work helps to label, describe, delimit and compare the concepts of a specific domain. It highlights the relations between these concepts and thus sketches a conceptual map of the domain.

b) Knowledge transfer

Terminology work allows explaining, comparing and conveying the concepts of a specific domain. In its multilingual form it helps finding equivalent concepts and related terms in other realities and languages.

Serving either of these functions or both, terminology work might be

- **monolingual**, when it is carried out in one language only
- **bilingual/multilingual**, when it considers two or more languages

- **prescriptive**, when the main aim consists in choosing, among different co-occurring variants and synonyms within a specialised domain, the most adequate or correct one, according to a set of predefined criteria, often with official standardisation in mind
- **descriptive**, when the main aim consists in collecting all competing synonyms and variants that are used to designate a concept, listing their peculiarities and usage limitations
- **translation-oriented,** when the main aim consists in supporting language mediators (i.e. translators, interpreters, etc.) in their daily work, sparing them the efforts of carrying out detailed and accurate research for equivalents and offering possible translation proposals where no existing equivalent can be found
- **multifunctional**, when it serves more than one function (e.g. it can be mainly descriptive, but with some prescriptive elements, or the other way round)

- **ad-hoc,** when it is based on a specific terminology problem or a limited set of issues that need to be solved for a specific task (e.g. translating a document)

- **systematic**, when it takes into consideration an entire domain or subdomain

- **text-based**, when it considers all terms contained in a specific text or set of texts

- **proactive**, when terminology is compiled before translation work, product development, etc. or any other purpose it serves

- **a posteriori**, when terminology is collected *after* translation work, product development, etc.

All these aspects may interact in various ways. For example, it is possible to do multilingual descriptive proactive terminology work or monolingual prescriptive systematic terminology work. The function and type of terminological activity influences the workflow, i.e. the way terminology is collected, organised, compiled and disseminated.

A particular value and importance of terminology work lies in its interdisciplinary approach and in its support to communication. In each domain different sets of specialised terminologies might coexist, according to the moment in time (diachronic variation), geographical area (diatopic variation), communicative situation (diaphasic variation), means (diamesic variation) or social situation (diastratic variation) in which they can or should be used. There are terms that are adequate for expert-to-expert communication, for expert-to-laypersons communication, for legislation, for academic writings, etc. In general, the more accurate this modulation of variation, the more efficient and effective the communication between people. Terminology collections help guiding the choice of the correct and most adequate terms in one or more languages according to conceptual equivalence, communicative situation, moment in time, etc.

Terminology can be considered an asset; it has an economic and legal value. For example, the terminology used and developed by a company is part of its know-how. Incorrect and incoherent use of terminology (in one or more languages) might reduce the quality of a product, cause additional costs and even lead to law suits in the worst case (product liability!).

2.2 Terminology work in the field of law

This is not the place for general introductions into terminology theory and work. However, given the focus of the LISE project on legal terminology and the peculiarities of terminology work carried out across different legal systems, we consider it useful to give a short overview of the main aspects to be considered when dealing with legal terminology.

Correctness and precision in legal texts are an essential prerequisite for legal certainty. It must be kept in mind that the law is actually *made of* language. Legal language, unlike other languages for special

purposes (LSP), not only serves to describe reality, but rather *creates* and *modifies* it. The law heavily influences society and the daily lives of every individual (e.g. buying a house, getting married and divorced, changing permanent address, etc.). As a consequence, legal texts need not only to be written in a legally correct and precise way, but also with great attention to the communicative aspect and to the efficient transmission of their contents.

In multilingual environments, where legal texts are drafted in several languages (e.g. the European Union, Switzerland, Belgium, etc.) maintaining precision, correctness and full equivalence across language versions is a demanding task. When language barriers are coupled with conceptual barriers, i.e. when the transfer takes place across languages *and* legal systems and so calls for legal comparison of legal concepts and cultures, conveying the message across these barriers becomes particularly daunting. Linguistic and cultural competences in the source and target languages need to be enhanced by the legal competences in the source and target legal systems. It thus becomes of paramount importance to devote particular care and attention to translating legal concepts and finding possible equivalents in the target legal system.

Correctness and precision in legal content and language can be guaranteed only when there is full awareness of the existing differences between legal systems. In fact, the diverging legal realities created within each legal system, the different taxonomies, are expressed by the legal language that is used to describe and convey them. Each legal language is therefore intimately connected with its legal system of reference. The national character of legal language is so strong that even countries sharing the same language often develop diverging legal terminologies and taxonomies (e.g. France and Belgium, Germany and Austria, Ireland and the United Kingdom, etc.). Every legal term must therefore be interpreted and understood within its legal system, background and culture, as it is embedded in a specific historical, social, cultural and legal context. This entails that translating a legal text is not limited to finding another linguistic label, i.e. in moving from one language to the other, but rather in moving from one legal reality to another. For example, there will be no real equivalent to the German term *Bundesstraße* (lit. “federal road”) in states that are not organised according to the federal model. Similarly, the Italian *certificazione antimafia* (lit. “anti-mafia certification”) will not correspond to any existing certification in most European countries.

At European level, notwithstanding the fact that there is one common European legal reality behind EU terminology, a challenge lies in the relation with the terminology and legal systems of the member states. It is commonly agreed on that EU terminology differs, sometimes substantially, from the terminology used at national level in each official language. Even when the linguistic labels as such might correspond, the concept they express can be slightly different, more precise or more generic than the one expressed by the national term. This again calls for extreme care and wariness when looking for equivalents and possible translations of a term. For example, it must be assessed whether the adoption of national terminology at EU level, but with a (slightly) different meaning, serves the principle of transparency and linguistic economy or whether it rather causes serious misunderstandings and risks hampering clear and efficient communication between the European Union and its member states.

3. Methodology

The research methodology applied mainly follows the qualitative approach. A careful analysis of the context, together with the experience of the LISE consortium members, allowed pinpointing the aspects that needed to be investigated in order to reach the project objective and deciding how to proceed. This led to the decision of interviewing a representative selection of organisations that follow recognised practices in terminology work. These selected terminology centres are located primarily in Europe, but also beyond. To collect additional data besides those gathered with the interviews, an online questionnaire survey was prepared and disseminated.

The results of both actions further described below allowed collecting useful input on the current practices in terminology work as well as their weak and strong points. Next to suggesting how the LISE tool suite and platform could be best applied within the different workflows, the evidence from the interviews and the responses to the questionnaire represent a solid basis also for drafting the guide for legal/administrative terminology work (deliverable D 3.3).

3.1 Semi-structured interviews

In the months between December 2011 and June 2012 two researchers of the LISE consortium have conducted 17 interviews with selected experts from organisations that carry out terminology work or manage terminology for various reasons and purposes. They had been chosen so as to obtain a representative selection of the main situations in which terminology is elaborated, covering a large number of domains, languages, objectives and methodologies (e.g. terminology for language planning, proactive terminology work to support translation, in private or public institutions, at local or international level, etc.). The sample therefore adequately mirrors and summarises the most important current approaches to terminology work. A particular focus has been laid on organisations that deal with legal terminology, at least to some extent. The interviews were conducted following a set of predefined open questions on the aspects to be investigated (semi-structured interviews), e.g. on terminology management, standardisation, etc., and all lasted between one and maximum two hours. The data gathered during the interviews represent the backbone of our research.

More in detail, the activities carried out during this phase were:

1. **Documentation and preparation**: detailed and in-depth bibliographic research and documentation on specific aspects of terminology elaboration, storage, management, dissemination, standardisation, common practices, etc., including user needs. This served as input so as to be able to draft precise and relevant questions during the interviews.

2. **Selection of possible interviewees**: 17 experts representing a good spectrum of situations in which terminology is elaborated and managed gave their consent for an interview.

3. **Definition of a semi-structured interview protocol**: the responsible partners prepared and discussed a detailed interview protocol with questions on general aspects, methodology, terminology management, terminology management systems and terminology planning (see attachment A). This ensured that all aspects relevant to each organisation and specific situation were always covered during the live interviews. At the same time, a semi-structured interview protocol allows for enough flexibility to investigate further aspects that might be touched upon during the conversation and that seem important for the LISE project. As a consequence of this, and also given the heterogeneous nature of the sample and the varied experiences, not all themes were dealt with in detail and not all questions were posed to all interviewees. All interviewees were first asked to talk freely about their activity and workflow. Then the interviewers posed specific questions to gather possibly interesting additional information. This approach put the interviewees at their ease, so that most conversations were even more informative than initially expected.

4. **Interviews**: Before each meeting the interviewers gathered available information on the specific organisations to be visited. The conversations were held with the people selected (see point 2 above) according to the predefined protocol (see point 3 above). All interviews were recorded, provided that the interviewees granted their consent to do so. Almost all of them were conducted face-to-face, one via conference call.

5. **Transcription**: Almost all interviews have been transcribed for a faster and safer collection of all relevant information.

3.2 Online questionnaire

Between mid-December 2012 and end of March 2013 an online questionnaire of 58 questions (see attachment B) was drafted and disseminated through the most important terminology networks and associations. It allowed collecting further 22 complete answers from organisations around the world, while limiting the travel and staff costs. Some of the questions in the online survey are very similar to the questions selected for the interviews, but most of them had to be formulated in a more detailed and precise way. Several questions also provide multiple choice forms listing possible answers to make compilation easier and quicker.

More in detail, the activities carried out during this phase were:

1. **Selection of sample:** A list of relevant networks and associations that could disseminate the questionnaire as widely as possible was drafted.

2. **Elaboration and adaptation of questions:** The interview protocol served as a guideline for the selection, adaptation, specification and detailing of the questions and the possible answers of

the online survey. Interdependencies and connections between questions were defined. The questionnaire was divided into four parts, focusing respectively on general aspects, methodology, terminology management and terminology planning.

3. **Survey**: The questionnaire was put online with the help of a specific programme that provides an online interface for compilation and stores responses. The link to the survey was then sent out to the sample of contacts and disseminators selected (see point 1 above).

3.3 Analysis and interpretation of data

The data collected with the help of the interviews and online questionnaire have been analysed to find the good practices, standard models and strategies currently adopted for terminology work, with a particular focus on the field of law. The collected data will help

- **assess the tools, approaches and practices** employed to do terminology work and manage the workflow, especially in the field of law
- **identify the needs** of terminologists and organisations doing terminology work, be they openly expressed or not
- **propose an approach** for answering the most frequent questions and solving common problems
- **picture the state of art** and, if necessary, propose how to modify, adjust and improve the standard approach to terminology work (aim of deliverable D 3.3; see point 3 below)

The approach followed foresees three consecutive steps:

1. **analysis:** collect, analyse and evaluate data with sociolinguistic methods
2. **interpretation:** identify meaningful and possibly informative relations, connections and interdependencies between the data
3. **application:** collect common problems and practical solutions, as well as open questions, priorities, needs or difficulties expressed, and interpret the data. Suggest possible answers and solutions to the most urgent and important current problems whenever possible. This also serves as a basis for deliverable D 3.3 (Guide for legal/administrative terminology work).

4. Report analysis of existing terminology workflows (D 3.1)

According to the specific needs and objectives of the organisation, terminology elaboration and standardisation workflows can be quite diverse, when looking at the details. For example, some standardising bodies might have an “express procedure” for urgent and important terms, others might not. Terminology teams may be composed of language and domain experts working together or the domain experts may be entrusted with checking the work previously produced by terminologists. Revision of each other’s work may be part of a terminologist’s daily chores or it might be the specific task of some senior staff member.

As a consequence, in the following analysis we have tried to abstract from single peculiarities and highlight the main tasks and activities that may commonly be carried out to produce terminology. These are not necessarily entrusted to different people and might sometimes be performed contemporaneously (e.g. term selection as documentation work proceeds). Also, not all organisations perform all tasks listed. For example, term extraction or the drawing of detailed concept systems are activities often skipped in some types of terminology work (see section 2.1), since the main aim might be to solve urgent needs (ad-hoc terminology work) or to collect all relevant terminology from draft legislation (text-oriented terminology work), for example. As far as dissemination and standardisation are concerned, there might be specific bodies entrusted with those steps within the workflow or they might not be performed at all. In fact, not many organisations actually officially standardise their terminology. Also, some keep their work in-house, as it merely serves internal needs.

Concerning official standardisation, we have decided to consider it in detail, notwithstanding the fact that a limited number of organisations are familiar with this type of activity, because the entire terminology workflow might be affected by the predominantly prescriptive approach. Also, terminology standardisation is one of the important aspects in the LISE project.

The two schemes below (Fig. 1 and Fig. 2) include the most common activities and put them in a clear order wherever possible, disregarding the fact that some steps might be skipped in practical daily activity. The more detailed bullet point descriptions in the subsequent pages include related activities as well as aspects to be considered, without going into the details of any specific workflow currently applied in any of the institutions considered by the LISE project. They are kept as short and compact as possible for maximum schematisation.

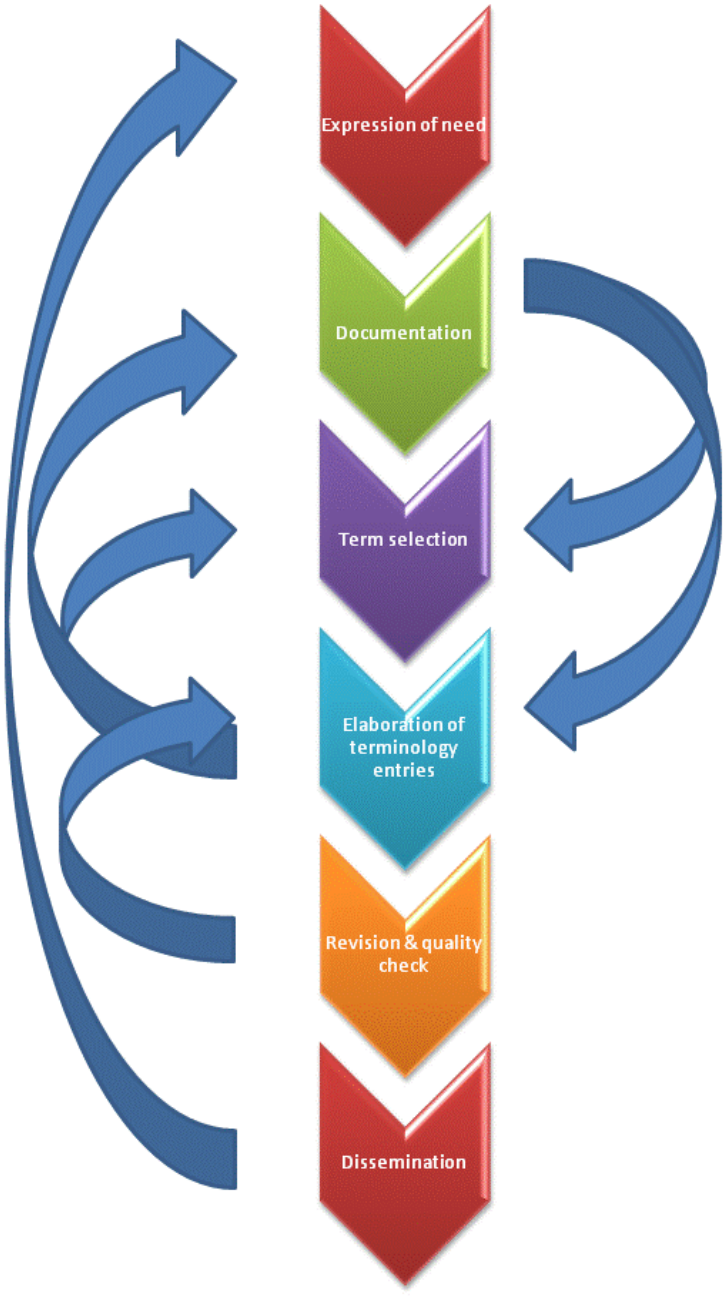

Fig. 1 Main activities of terminology work

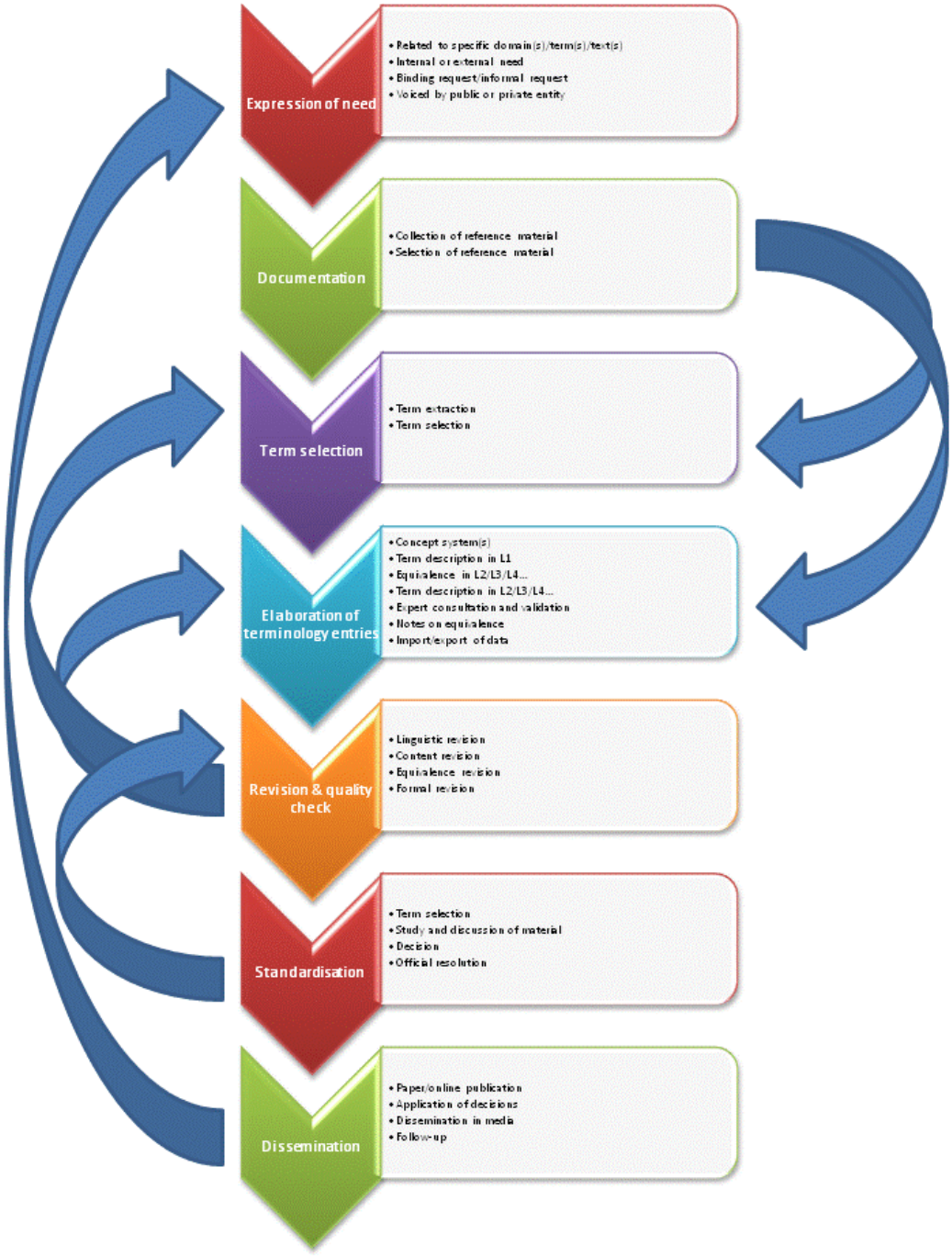

Fig. 2 Complete terminology workflow with standardisation

4.1 Expression of need

Related to

- specific domain(s)
- text(s) (e.g. to be translated)
- single term or group of terms

Voiced by

- internal entities
- external entities
- public entity
  - official translation service
  - any public organisation (government, public administration, court of justice, university, etc.)
- private entity
  - private organisation (e.g. professional associations)
  - private citizen

Expressed via

- formal binding request
- informal request
- official feedback and request form (e.g. macro for suggesting terms)
- informal channel (e.g. email, telephone, Word lists, etc.)

N.B. In order to respond to the different types of needs expressed, terminology work might be carried out systematically, in a strongly text-oriented way or ad-hoc (see 2.1). Proactive terminology work aims at anticipating the expression of need as much as possible.

4.2 Documentation

Collection of documentation and reference material

Selection of documentation and reference material

N.B. Reference material can be of extremely diverse nature, internal or public, printed or on other media (esp. the Internet), etc. Often these sources have a precise hierarchy of relevance and importance, defined by the (group of) terminologists according to the project that is being worked on. In legal terminology work, for example, legal codices might be considered more relevant than handbooks or sentences, which in turn will be more important than Ministry websites, etc.

4.3 Term extraction

Procedure

- manual extraction
- (semi)automatic extraction and manual check

N.B. There are many available tools for automatic term extraction. Large amounts of raw or pre-processed text are needed to use them. Also bilingual term extraction from aligned parallel texts (e.g. translation memories, aligned multilingual corpora) is possible. Whether these tools are employed or not might depend on the available text material, the languages treated, the type of terminology work (systematic or text-based, not ad-hoc), the size of the project, the familiarity with the tools, etc.

4.4 Term selection

Creation of list(s) of terms to be elaborated on the basis of:

- systematic approach (e.g. concept tree)
- ad-hoc (e.g. specific needs expressed)
- selected texts (text-based approach)

Approval of term selection (e.g. by project manager, domain expert(s), contractor, etc.)

N.B. When terms are extracted manually, a pre-selection is usually performed during extraction. Term selection is one of the steps where domain experts might be involved more systematically.

4.5 Elaboration of terminological entries

Elaboration of concept systems

- ordering concepts according to a conceptual classification (e.g. thematic scheme)
- structuring the concepts in a clear and systematic scheme
- identifying the relationships between the selected concepts
- adapting conceptual systems to results of comparative work after term description in L1 (source language) and L2/L3/L4 (target languages) etc. (see below)

Term description in L1

- finding/writing definitions, contexts of use, grammatical indications, etc.
- retrieval/selection of synonyms/variants

Equivalence in one or more languages

- manual looking for equivalents
- (semi)automatic extraction from parallel aligned corpora/translation memories

Term description in L2/L3/L4 etc.

- finding/writing definitions, contexts of use, grammatical indications, etc.
- retrieval/selection of synonyms/variants
- translation proposals in L2/L3/L4 etc. to fill terminology gaps

Notes

- on equivalence (e.g. partial equivalence, etc.)
- on peculiar usages (e.g. limitations in time, register, text typology, etc.)
- on direction of translation

Expert consultation and validation of

- terms/terminological entries
- conceptual equivalence

Import/export of external data

Harmonisation of imported data and legacy data

N.B. Concept systems serve to model concept structures on the basis of the specialised knowledge of a domain, to clarify the relations between concepts, to create the basis for a uniform (and standardised) terminology, to facilitate contrastive analysis of concepts and designations across languages and finally to help writing definitions. Nevertheless, the formal elaboration of concept systems is often skipped, especially in some types of terminology work. It is typical of systematic domain-based terminology work. The level of detail of the terminological entries that are elaborated can vary depending on the type of terminology work (e.g. the presence of usage notes might be extremely important in translation-oriented terminology work, definitions might be considered essential in standardisation-oriented terminology work, etc.). Concerning domain experts, these are often asked to check and revise monolingual work, more rarely are they involved in validating conceptual equivalence.

4.6 Revision & quality assessment

Overall quality assessment (QA)

- checking equivalence between concepts expressed by terms in all languages
- content revision (e.g. correctness of definitions, etc.)
- language revision (e.g. linguistic correctness of definitions, etc.)
- formal revision (e.g. presence of all data categories, correct domain attributions, etc.)

- periodic/regular revision and QA
- on-going revision and QA

Validation

- by (senior) terminologists
- by domain experts

N.B. Revision predominantly takes place on paper (e.g. printouts of data exports). One of the additional steps performed by some terminology groups at the beginning of any project is the selection of (multilingual) domain experts for revision. Quality assessment comprises all those actions that are necessary to ensure the correspondence of a terminological entry as well as an entire collection to predefined criteria of correctness, completeness and coherence (e.g. adherence to formal rules). It ensures that the work responds to the needs of the end-users. When duplicate entries (doublettes) are spotted during QA, consolidation and harmonisation steps are triggered so as to determine which data shall be merged, deleted or further worked on.

4.7 Standardisation

Selection of term(s)/domain(s) to be standardised in one or more languages

- systematic approach (e.g. concept tree)
- ad-hoc (e.g. specific needs expressed)

Study of proposed terms/equivalences/translation proposals/specific reference material (e.g. complete terminological entries)

Discussion taking into consideration

- specific criteria formally defined
- no formalised list of criteria
- in physical meetings
- via written procedure
- in forums, mailing lists, etc.

Decision (e.g. eliminating or reducing synonyms in both languages, determining 1:1 translations, validation of translation proposals)

- in physical meetings
- via written procedure
- in forum, mailing lists, etc.

Revision of concept descriptions (e.g. definitions)

Official resolution (decree, resolution, drafting of minutes, etc.)

- of terms in one language
- equivalents

N.B. Standardisation committees are usually composed of language experts (especially when dealing with lesser used languages) or terminologists/translators and domain experts. The latter usually have a good (passive) command of all languages treated, next to their domain-related expertise. Standardisation committees generally work on the basis of detailed research by the terminologists and predominantly take decisions during physical meeting. Non problematic terminology previously circulated via email or on paper is validated without further discussion during the meetings. Some standardisation committees foresee the possibility of re-discussing decisions taken in the past when a social, legislative, cultural change, etc. calls for a revision and new standardisation. Some committees might be subdivided in several domain-related sub-committees. Their decisions are published online, on printed media (e.g. dictionaries) or in Official Journals. In business environments the group of standardisers is often called the “terminology circle”.

4.8 Dissemination

Publication of work in

- terminology databases
- dictionaries, glossaries, etc.
- official documents (e.g. Official Gazette, etc.)

Direct use in

- official/legally binding documentation
- all realms of public and private life

Dissemination in press, radio, TV, etc., esp. of standardisation work

Follow-up (e.g. checking success in implanting new terms)

N.B. This seems to be one of the most delicate steps of the terminology workflow. Making data available to the intended end users and/or the public at large, allows collecting external feedback. Terminology

databases contribute to greater transparency in the domains treated, may favour multilingualism and support equal access to information. Next to disseminating data through dedicated termbases or publications and publishing the results of standardisation also in dictionaries, Official Journals etc., terminology needs to be disseminated widely in order to make sure it is actually used and implanted, even more so if it is officially standardised. Some terminology centres provide periodic newsletters, brochures, information days, trainings to newly employed staff, intranet blogs, etc. Also dissemination on the general media (newspaper articles, radio and TV broadcasts) is useful to inform about the work done, its purpose and usefulness.

5. Consolidated analysis of data

The relevant but still limited number of responses to the questionnaire and the interviews do not allow a large-scale quantitative analysis. We shall rather focus on the qualitative aspects, thus reporting on the most significant data and highlighting tendencies, common practices, interesting aspects, shared attitudes, desires frequently voiced and basic problems that emerged from both sets of data. The interpretation also takes into consideration the experiences of the LISE consortium partners and the information gathered during the study of the published material and grey literature (e.g. internal guidelines, documentation, workflow charts, etc.). We have tried to postulate motivations where these were not made explicit during interviews as well as to illustrate the background and consequences of specific choices and practices.

It must be reminded that, while the interviews were mainly conducted with terminology centres in public or semi-public organisations that predominantly (directly or indirectly) support translation services and language centres, the questionnaire was open to all kinds of public and private organisations doing terminology work. As the responses to the online survey were anonymous we cannot tell how many private companies and businesses participated in the research. However, answers to some questions (e.g. Who is terminology work done for? Is it freely accessible?) seem to suggest that the public and semi-public organisations prevail also in this part of the research.

The following analysis is subdivided into four sections, i.e. general aspects, methodology, terminology management and terminology planning.

5.1 General aspects

More than three quarters of the organisations involved in this research have been producing and managing terminology for over ten years. The very first organisations started collecting terminology and employing people for that purpose about half a century ago, but there are also very few that have only recently decided to elaborate terminology in a more systematic and officialised way. This indicates the high level of expertise and in-depth experience with terminology work of all organisations that conferred information to the project.

People doing terminology work are almost exclusively language mediators, i.e. translators and interpreters, or terminologists proper. Some institutions also have a limited number of domain experts directly working on the elaboration of terminological entries (see below on the main role of domain experts). As a consequence, all staff in terminology work holds a university degree, mainly in translation

or interpreting, but also in other language-related disciplines (e.g. modern languages), with the exception of trainees who are often still completing their studies. Due to the limited amount of specific university courses, rarely people hold a degree in terminology. It is quite common to have acquired the methods of terminology work by doing it, possibly with the help of senior supervisors, or in specific training courses (e.g. Terminology Summer School and similar initiatives). This is particularly true for senior staff members. The data suggest a specific need for more training courses in terminology work and for a greater focus on this type of activity at university level. One of the problems voiced in different occasions is the lack of qualified well-trained job candidates on the market.

The native speaker principle is generally not strictly applied in about two thirds of the organisations considered, i.e. it is not demanded that each staff member mainly/exclusively work on a specific language (e.g. editing only one language section in all terminological entries). Only a very limited number of organisations can afford this, as it implies notable investments in human resources. Some organisations do, however, require all entries to be at least reviewed by native speakers of each working language. The interviews underline the importance of passive language knowledge. It is indeed not rare for terminologists to actively try to learn new languages and improve their (passive) language competences in training courses, which may partly be financed by their employer.

Terminology work is first and foremost aimed at supporting language mediators, but also domain experts. It mostly serves internal purposes. The general public is a further (sometimes secondary) target group. The next end user category is public/private organisations. A limited number of terminology centres caters for customers. This figure probably indicates that, as with the interviews, most respondents of the online survey are public institutions rather than private companies. The latter obviously provide terminological data collections and expert advice exclusively for their clients.

More in detail, the main categories of end users directly targeted at, are therefore:

- translators/interpreters of in-house language mediation services
- translators/interpreters of language mediation services in related/affiliated organisations
- freelance translators/interpreters working with the organisation
- other in-house staff that do not belong to the translation/language services but need and use the terminology

With rare exceptions, most terminology centres consider themselves short of staff and would wish for more personnel. The average size clusters around two to five full time staff, with one large organisation employing a maximum of forty staff members. This indicates that there is a growing demand for terminology based on real needs. Also, that terminology work is considered an intellectual human activity. As a matter of fact, it might be surprising to realise that the tools that might *replace* work done by humans are not commonly used at all (basically term extractors only), whereas all organisations use

tools that *support* their work, even though to different degrees of satisfaction (mainly terminology management tools) (see section 5.3 below).

In organisations where the terminologists predominantly direct and coordinate work and which maintain a large translation service, the principle of translator secondment to the terminology service is a way of

- enlarging the number of available staff
- teaching translators the principles and rules of terminology work
- collecting input on real needs
- fine-tuning terminology work to respond to the daily needs of translators

The number of translators on secondment ("rotating terminologists") may vary, according to the size of the organisation, between one and four at a time.

Most terminology services also have team assistants for basic administrative work. Very few larger ones have internal or dedicated technical assistants. This is obviously the case when in-house programmed software is used for terminology management (see section 5.4).

Traineeships are common and welcome. They ensure enthusiastic staff that is willing to learn and contribute with quality work. Trainees keep the contact with the academic world and help disseminating the importance and methods of terminology work. They promote the use of the organisation's own terminology database (where present). Traineeships also allow spotting possible highly qualified future staff members, if desired.

5.2 Methodology

The descriptive and translation-oriented approach generally prevail over the prescriptive, since a very limited number of organisations consider terminology standardisation proper as their main aim, while many organisations support translation services. Nevertheless, different approaches tend to co-exist: standardisation may not be a prevailing objective, but a purely descriptive attitude would probably fail in fostering a better and more coherent communication. As a consequence, the work of most terminology centres has a descriptive character, but a certain amount of preliminary selection (e.g. through research or consultation with experts) often introduces a prescriptive element. For example, one institution clearly specifies that it generally follows a descriptive approach, but sometimes considers it necessary to be prescriptive, even without the help of an official standardisation committee or board. Another one similarly works in a descriptive way, but then hands over the terminology to experts of its members. These then decide which terms should be used following a prescriptive approach.

Within terminology centres serving as a support for translation services terminology work is naturally translation-oriented ("terminology helpdesk") or multifunctional. Terminology work in that context might therefore be considered either aimed at supporting text production, translation or, less often, standardisation. These terminology centres predominantly solve specific terminological problems (ad-hoc terminology or also "express terminology", when the answer to a terminological issue is extremely urgent). One organisation estimates as much as 90% ad-hoc terminology work and 10% systematic work, due to lack of personnel and time constraints. This was called doing "day-by-day terminology".

Even though few organisations declare to aim at official terminology standardisation, standardisation-oriented terminology work seems to be less rare than expected, for the reasons explained above concerning the implicit prescriptive element. Also, it is probably more common in private companies where the main aim is to collect a coherent and consolidated set of terminology for internal and public use. Similar considerations hold true for systematic terminology work, where an entire domain is covered and elaborated in a coherent project. It is considered time-consuming work and would call for more staff. Few organisations therefore exclusively work systematically, while many actually do keep it as an on-going parallel activity when there is enough time.

What is gaining more and more importance is the so-called "proactive" terminology work. On the basis of future terminological needs determined together with the target groups or user groups (e.g. with the text drafters or the translation service), many terminology centres now try to cater in advance for future needs. For example, they try to find out which themes will be put up for discussion in the following months, which laws are to be passed, etc. After having cross-checked with what is already available in their databases or in other resources, they elaborate the missing terminology in the languages needed. This is definitely a positive and very welcome development in terminology.

In general, terminological support should be available even before it is actually called for, especially in a translation environment. It then accompanies the entire translation, text production or product development process whenever needed. Not many organisations actually take care of collecting, describing and managing terminology a posteriori.

Law and administration, followed by science and technology are the domains that prevail largely over all the others (e.g. economy and finance, medicine, information technology, etc.). Particularly important for those that deal with legal terminology is administrative terminology. Next to supporting mainly administration, some organisations also cater for the judiciary and legislative powers.

Terminological comparison in the realm of law is rare. Most terminology centres that treat legal and administrative issues deal with one legal system, although in several languages. This is because they support their national or local bodies and translation services. As a consequence, the methodology applied to any other domain is used also for elaborating legal terminology, with a few rare exceptions that actually made legal comparison their core activity. The limited number of institutions that

introduced at least some comparative elements in their work usually specify the geographical usage, i.e. the legal system of reference, as well as other usage restrictions for every term treated in the database.

The core legal domain of the LISE project, social security, has been taken into consideration by less than half of all organisations that work with legal terminology. This proves that there is a need for more terminological comparison and terminology work in this specific subject field.

More than half of the organisations involved in the research plan to add further domains to their terminology databases in the future. The following options and criteria are listed

- new domains that might be relevant for their activity
- new domains will definitely be added, but it is not yet clear which ones
- all domains and subject fields regulated by new legislation
- all terminology that is somehow related to their new products
- some specific domains of interest, e.g. social security, consumer law, international criminal law and medicine

The number of working languages is an average of four. Few organisations work on 23, 27, 30, even more than 40 languages. The focus of the LISE research being Europe, it is not surprising that English and French are almost always one of the working languages. German follows, but is used slightly less often. Italian and Spanish, then Dutch and Portuguese come next. All other European languages are considered in a more limited number of organisations. Lesser used languages like Basque, Galician and Welsh close the list. Interestingly, one organisation uses one major European language not as working language but as a point of reference for the elaboration of translation proposals in its main language. About half the organisations do not exclude the possibility of adding new languages in the future. The others are almost equally divided between those who positively plan to do so and others that already cover the languages they need. In general, the languages that will be added are Asian languages and other languages spoken in fast growing developing countries.

The working languages usually correspond to the languages spoken in the country where the organisation is located, including in some cases the official minority languages. This is true especially for those dealing with legal terminology. One public organisation has included the languages of present commercial partners and is considering adding a further language spoken in a large country which is of growing commercial importance.

As explained in section 2.2, legal terminology is strongly connected with the social and cultural reality in which it was created. As a consequence, a common problem encountered when trying to translate a legal term is that no equivalent concept can be found in the legal system of reference of the target

language and therefore, no term either. These terminological gaps need to be (temporarily) filled in some way in order to ensure efficient and clear communication across languages and legal systems. Terminological issues are solved almost exclusively by terminologists or domain experts, either together or separately. When the issue is multilingual, it is again a matter for terminologists, then domain experts and finally also for translators. Sometimes technical writers or the customers are involved in solving both monolingual and multilingual terminology issues. However, neologisms are clearly reserved for terminologists and domain experts: they are created predominantly by terminologists or domain experts, either together or separately. More rarely, translators and other people, such as technical writers or customers, may be involved in the process. Validating the neologisms is principally a domain expert's task, unless there is a dedicated committee or board. More rarely, it is the terminologist, or even less often the translator who validates terminology.

Some formalised procedures for the creation of neologisms are:

- Neologisms are (proposed and) validated by specific boards or committees, composed of domain experts and language experts (for example legal experts and jurilinguists), by translators and domain experts or by linguists (especially when dealing with minority languages). These committees often base their work on a specific set of criteria and use preliminary work prepared by terminologists as a basis. Once the neologisms have been approved, they become official and should be used in all public and official contexts. This is the common procedure in countries where terminology is officially standardised.

- Neologisms are created by specific offices and boards (e.g. directorates, departments, etc.). The terminology centres are consulted for help and advice, but the last decision remains with the office or board.

- Neologisms are created by the language services of each member state, as these might have organisations in charge of language development and safeguarding. The member states are free in selecting and proposing the necessary neologisms, but the final decision remains with the terminologist within the terminology centre.

- The translators collect material and study documents in order to propose adequate neologisms, which might be reviewed by domain experts and native speakers, if the translators consider it necessary. The final decision remains with the terminologist within the terminology centre.

- Neologisms are proposed by terminologists or by terminology working groups. The proposal is reviewed by translators who are native speakers or, more often, by the coordinator of the relevant language department. Finally, the proposal is discussed with the customer, who takes the final decision.

- Neologisms are created only very rarely, with the exception of the lesser used languages treated. In that case other languages from the same language family might be used as a point of reference.
- Neologisms are avoided. Instead, explanatory paraphrases of the concept to be conveyed are discussed with domain experts (e.g. jurilinguists) and terminologists.

The role of domain experts is particularly important when creating new terms. The data show that, unless an official validation or standardisation board exists, domain experts are very often involved with an informal procedure, on the basis of goodwill and of the terminologists' personal contacts. It is quite rare to encounter organisations whose in-house domain experts are formally and regularly involved in the terminology workflow. This is obviously due to the fact that the domains treated might be diverse and not available in-house. Nevertheless, terminologists do declare to consult experts regularly, whenever needed. Every terminologist has his or her own network of contacts to be asked for advice mainly via phone or email in an absolutely non-formalised procedure for validating terminology. The option of setting up a more regular and constant consultation with domain experts is an often voiced desire of terminologists. Also, phone calls and email exchanges are an easy and quick way of communicating, but a common discussion forum (e.g. on the intranet) and a safer way of keeping track of discussions and proposals would be another welcome feature.

The reference material used for terminology consists in handbooks, manuals and similar material (in about three quarters of the organisations). Legislation and technical reports & documentation follow *ex aequo*, probably depending on the domains treated. In two thirds of the organisations internal documentation is popular reference material. Slightly more than one in two organisations also exploits the World Wide Web. A minority refers to databases, translation memories and standards. Again, this depends very much on the availability of a specific type of reference material in the domains treated. The official sources like legislation, standards and the traditional printed publications still remain the main points of reference, but the World Wide Web is of growing importance, especially in some domains that are undergoing rapid development.

The use of corpora is quite diverse. No terminology centre declares to *always* employ corpora. About a quarter of the organisations involved never use them at all, while the remaining share uses corpora to a varying degree. Almost all organisations that work with corpora have internal ones and about two thirds also refer to freely available text collections. Both internal and freely available corpora are to a large extent bilingual or multilingual (about nine in ten) and aligned (two out of three corpora). Monolingual corpora are less popular but still considered useful. The use of corpora is more common in large-scale projects and when a huge amount of text has to be processed in a very short time. The same holds true of term extractors (see 5.3).

5.3 Terminology management

Terminology management is centralised in almost two thirds of the cases, but in large organisations there are often several departments collecting their own terminology. Data storage on the contrary is almost always centralised. Terminology management is a task for terminology managers, terminologists or senior terminologists. Together, these three groups cover about two thirds of the responses, with the terminology managers proper being about a fourth of the total. It must be added, though, that few organisations have a clear-cut definition of roles and tasks, some actually distinguish none at all. The two most common distinctions are between terminologist and translator/interpreter. Next to these two roles, the terminology managers might be present in a more limited number of organisations. About a sixth of the organisations also have the role of senior terminologist and of domain expert clearly attributed to their staff.

About three quarters of the organisations have one or more standard workflows for terminology production. An ever higher number, about four fifths, has written or uses specific guidelines for some or all aspects related to terminology work (e.g. data input rules, style guides, etc.). These are either for internal staff only (in about half of the cases) or only partly available to the public. Despite the presence of written guidelines, the importance of direct training of new staff members or trainees is never underestimated, on the contrary: personal coaching ensures a possibly homogeneous approach to terminology work within an organisation and enhances quality.

While the guidelines are considered of limited interest for the external world, terminological data are usually made public via dedicated web services and portals. Only one quarter of the organisations limits the access to their data to internal staff and to some selected (affiliated) organisations/people. This trend is actually changing, since various institutions decided to offer general public access to their terminological data collections in recent years. It suggests the desire of making data available to anyone who might need it, so as to maximise the effects of the human and financial resources invested, and is probably more typical of large public or semi-public organisations than of private companies. One organisation even hands over lists of standardised terms in specific training courses.

A common feature of terminology databases is a layered, often double, level of access, where internal staff (or the terminology group only) has access to more or more detailed information than the end users at large (about two thirds of the organisations). Generally the finalised or validated data are published, while the work in progress is kept for internal use only, at least until it can be released after validation.

At the beginning of the terminology workflow there is one of the most important activities in terminology work, i.e. term extraction. More than half of the organisations tend to perform this task manually and a small percentage actually never needs to do any extraction at all. These might receive their terminology directly from the translation services or from their customers. Semi-automatic term extraction takes place in about a third of the organisations. In general, term extraction is obviously less relevant for institutions doing ad-hoc terminology under great time constraints as a support to translation services. However, it is surprisingly limited also in environments where translation memories (and as a consequence, parallel aligned texts) are available and in standardisation-oriented organisations where systematic domain-related terminology work prevails. The quality of input provided by automatic term extractors is still considered unsatisfactory. Using this type of software is sometimes seen as a "loss of time". Furthermore, the currently available (commercial) tools are not always considered sufficiently easy to use. Finally, especially for the smaller centres, manual work is more immediate and more complete, also because the reading (or text skimming) process that accompanies manual term extraction provides at the same time understanding of the contents of the domain or text that are being treated. Experienced terminologists spot term relations, particular contexts of use and possible peculiarities of the terms while reading, placing them in a real or mental draft concept system. The use of term extractors is more common in large-scale projects and when a huge amount of text has to be processed in a very short time. The same holds true of corpora (see 5.2).

A more or less standardised terminology workflow is usually triggered by:

- lists of terms that must be checked, selected, updated and further researched upon
- texts that are handed over regularly (e.g. once a month) which contain useful terminology to be extracted (manually)
- the analysis of a specific situation or domain followed by terminological research
- regular meetings with (in-house) translators who voice their current and future terminological needs
- regular meetings held to assess which issues will be put up for discussion in the future (e.g. which legislation is likely to be passed), which domains may be particularly relevant for the organisation or the public, which domains have not been processed or need to be updated
- specific requests from translators, other internal staff, customers, the public, etc.

The terminological entries are then elaborated by terminologists who add any information specified in the data categories and required by the organisation's guidelines, where available. Usually the information provided is: definition, context, notes, sources, geographical usage (especially in legal terminology), synonyms, variants, equivalents, etc. The vast majority of terminology databases try to provide their users with a definition of the concept, wherever possible. Sources are always indicated according to internal rules and help users understand where the data come from, their relevance and time frame.

Another important workflow steps in terminology is revision. The staff who revises terminology work first and foremost consists in domain experts, closely followed by terminologists and senior terminologists. These two last categories add up to almost half of the revision work done. Most revisions are therefore done with a terminologist's eye. In many institutions revision has to be taken care of by native speakers. Sometimes translators/interpreters act as reviewers and in rare cases it might be technical writers or dedicated committees and boards.

Terminologists first and then senior terminologists are entrusted with quality assessment: more than a third of the work is done by these two categories. Also domain experts take care of checking the quality of the work in about a sixth of the organisations, followed by the terminology managers. Other professionals, such as translators/interpreters, technical writers or language experts (the latter dealing in particular with lesser used languages) are involved more rarely. As a general rule, revision seems to lie often with the domain experts, whereas terminology experts are called to control quality. The domain experts are therefore asked to revise the content of the entries, while the formal (correctness, coherence, consistency, etc.) and linguistic (grammar, orthography, etc.) checks are more naturally expected from the terminology and language experts.

Terminological entries are always updated and corrected manually. In about a third of the cases also batch changes and automatic updates are performed.

Termbase definitions are almost exclusively prepared by terminologists or senior terminologists. Together, these two categories take care of selecting and organising data categories in almost three quarters of the organisations.

Terminologists are usually also in charge of data imports and exports: about half of the responses refer to them. IT experts or computational linguists follow, with a slightly higher percentage than the terminology managers. It seems that IT people are involved mainly for large sets of data, while terminologists are usually entrusted with all other minor imports.

A good and efficient terminology management, especially for significant amounts of data, usually needs the support of dedicated tools. Three quarters of the organisations use such tools; basically all of them have developed or acquired a terminology management system. As we have seen, many organisations support translation work. As a consequence, almost two thirds of them also use translation memories, especially the concordancer, which allows finding equivalents. This explains how and why many organisations elaborate terminology during the entire translation process. It should be considered that translation memories generally allow some direct interaction with terminology management systems, that is, translation pairs can be stored in the terminology database both during translation and also at the end of the process. This makes the collection of raw terminology data easier and quicker. Terminology extractors score about three (out of ten), as do translation workflow tools. One in four also uses controlled language checkers. Finally, authoring memory tools are rarely used (about one in ten).

The tools developed in-house slightly prevail over the commercial tools, being used in one out of two cases. The commercial tools (over a fourth) are sometimes adapted to specific needs. If the share of commercial off-the-shelf tools is added to the share of adapted commercial tools, they total a slightly minor amount than in-house tools. Several organisations preferred an in-house developed tool for three main reasons:

- having started with terminology work many years ago, the commercial software available at that time was not considered adequate to meet their needs
- the in-house solution allows for maximum flexibility and independence
- the constant support of IT personnel ensures a regular implementation of all necessary updates, amendments and improvements to the data and their structure

It is surprising how many organisations using either, in-house or commercial tools, rely for example on Excel sheets and Word files to exchange and discuss terminology. This suggests that at the moment neither type of tool sufficiently or adequately supports information exchange, discussion, data exchange and collaborative work on terminology.

New languages are usually added manually and selectively to existing terminology collections. Only a small number of organisations actually do have experience with automatic processing and batch addition of data. One in five has no experience with adding languages at all.

Data are merged semi-automatically. Only a very small percentage of organisations perform this task manually and for rather small sets of data. Similarly, only in very limited occasions are data merged fully automatically. It quite likely seems too risky. One in five has no experience with data mergers at all. To be able to merge databases and sets of data, an interchange format is needed, such as TBX. About half of the organisations use such standard formats.

One of the consequences of data mergers might be the creation of doublettes, i.e. duplicate entries. These can also be created during manual input and update over many years. Doublettes pose a problem in all terminology databases: they confuse the end users and are not always easy to spot. The cleaning process is usually very much manual human work. Some organisations foresee periodic checks and systematic clean-ups, some take the occasion of checking their data for doublettes whenever the terminology database is updated, other consider it part of every-day work. In the latter case, doublettes are removed or merged whenever a terminologist stumbles upon them. The latter is particularly common in institutions that own large collections of data and do not have sufficient IT support or human resources to perform systematic clean-ups.

5.4 Terminology planning

As explained before, even in organisations where terminology work is not exclusively aimed at official standardisation, some prescriptive elements are always present. In the majority of organisations that do some standardisation work there are specific guidelines. The criteria for standardisation are set equally by terminologists or by standardisers (standardising committees, language boards, dedicated councils, etc.). Actually it is more common for terminologists to take care of terminology standardisation (one in two organisations), while few have a dedicated board or committee. Also domain experts tend to be involved in standardisation, but to a more limited degree than the terminologists (about one in four). Very often (regular) physical meetings take place to discuss terminological issues and collect solution proposals. Online meetings and dedicated forum discussions are limited.

In general, it seems that standardisation work needs much more software support. The desire for closed discussion forums, automatic distribution of lists of terms to be discussed and an automated way of tracing discussions or collecting decisions would be absolutely necessary to speed up work and enhance its quality.

5.5 Main issues addressed

The issues addressed by the interviewees and partly emerging from the answers to the questionnaire are:

- shortage of staff

  This is mainly due to budget limitations and the growing cuts in human resources.

- limited number of available qualified staff

This is partly due to the fact that there are not enough university courses and specialised trainings for terminologists.

- notable time pressure that often does not allow systematic work and reduces the time available for in-depth research to a minimum

This could be partly solved by employing more staff, but the access to material, documents, tools and experts should be enhanced, too.

- limited support provided by the available tools; the amount of information shared and discussed in Word documents and stored in Excel files (e.g. before batch import) is significant and to a certain extent surprising

Terminology management systems are mainly used for data storage, but many other activities, such as information collection, expert consultation, discussion, even revision take place outside the terminology management system.

- insufficient flexibility/adaptability of commercial tools, which lead various organisations to develop add-ons or alternative software in-house

This is especially true for organisations which set up their terminology services many years ago, when there were not many good or adequate commercial tools available.

- difficulty of involving (multilingual) domain experts

The consultation of experts too often relies on personal contacts and on the goodwill and spare time of the experts.

- difficulty of explaining domain experts how they are supposed to check and revise or comment on the data

Domain experts change according to the domain treated, many have never seen a terminological entry and it is up to the terminologist to explain the principles of terminology work, how revision should be done, etc.

- difficulty of collaborating with other institutions doing similar work and exchanging data (administrative and technical problems, copyright restrictions, etc.)

There is a clear reluctance in exchanging data. Data exchange tends to take more time than expected in preparation and needs to be well checked after completion. Also, there is the risk of losing or corrupting data. If problems with copyright or inter-organisational agreements arise, too, most terminology centres give up on the idea of merging or exchanging batches of data.

- lack of adequate reference material and documentation

In some new and quickly developing domains as well as in some lesser resourced languages the available documentation is simply not sufficient to produce (enough) high-quality terminological entries.

- difficulty of disseminating their work

Many terminology centres strive to make their work more widely known, convey its importance to everyone, and ensure the constant use of the terminology database. Some terminology centres within large organisations try to keep interest high and discussion lively by writing periodic newsletters or intranet posts. Still, the results are not always as positive as expected.

6. General scenarios in terminology workflow

The following scenarios illustrate typical situations that owners of terminological data collections might be facing. They describe the typical actions that can solve an existing problem or that are needed to perform a desired step. Each scenario presents also the adaptation of these workflows to the use of the LISE tools.

Scenario 1: Terminology database to be updated

Institution ONE has a large terminology database that has grown over the years. It covers different domains and languages, has been worked on by many people and is translation-oriented. One of the languages treated in it has undergone an official spelling reform.

Solution	Description	Activities and consequences	Advantages	Disadvantages
1	Do nothing.	End-users will take care of applying the reform themselves.	No additional efforts concerning person/months and financial resources.	No control over correctness of spelling. Outdated spelling may confuse and mislead end-users, thus decreasing the trust into the terminology resource.
2	Slowly change entries that contain the reformed	End-users will take care of applying the reform themselves for some	Limited additional efforts concerning	Long-term activity that risks being trailed along

Solution	Description	Activities and consequences	Advantages	Disadvantages
	language manually, as they happen to be worked on (also for other reasons).	entries but not for all.	person/months and financial resources.	for months/years. Difficult to determine when activity will be completed and no overview over work already done. The termbase will contain a mix of terms with new and outdated spelling for a possibly endless amount of time. Presence of outdated next to reformed spelling may confuse and mislead end-users, thus decreasing the trust into the terminology resource.
3	Systematically change all entries that contain the reformed language manually.	End-users will take care of applying the reform themselves for some entries but not for all.	Depending on the extent of the reform, limited additional efforts concerning	Depending on the extent of the reform, notable additional efforts concerning

Solution	Description	Activities and consequences	Advantages	Disadvantages
			person/months and financial resources. End of activity can be foreseen and planned.	person/months and financial resources. The termbase will contain a mix of terms with new and outdated spelling for a limited amount of time. Presence of outdated next to reformed spelling may confuse and mislead end-users.
4	Batch change and update all entries that contain the reformed language.	Text fields in the reformed language or the entire terminology database are exported, automatically processed with the help of an ad-hoc or available script/programme, then imported back. End-users will access a terminology database with new correct spelling immediately after batch change.	Quick and clean. Depending on the extent of the reform and the availability of scripts/programmes, limited additional efforts concerning person/months and financial resources.	Depending on the extent of the reform and the availability of scripts/programmes, notable additional efforts concerning person/months and financial resources. Needs qualified staff able to process data (import/export), write

Solution	Description	Activities and consequences	Advantages	Disadvantages
				and run the script or use the programme. Purchase/retrieval of programme or writing of script.
LISE	Batch change and update all entries that contain the reformed language.	Text fields in the reformed language or the entire terminology database are exported, automatically processed by ESTeam, viewed and approved/rejected by own terminologists, then imported back by ESTeam (CLEANUP). End-users will access a terminology database with new correct spelling immediately after batch change. If other CLEANUP routines are implemented, the terminological data collection will be cleaner, too.	Quick and clean. Limited additional efforts concerning person/months. Terminologists can view, approve or reject every single change proposed by the LISE tool CLEANUP, if desired. Possibility of discussing any change within the LISE service platform in real time. The CLEANUP tool allows performing further improvements (remove	Purchase of LISE tool suite.

Solution	Description	Activities and consequences	Advantages	Disadvantages
			typos, comply to canonical forms, etc.), if desired.	

A common case in the business world is companies changing language region and thus choosing to replace British English with American English, for example.

The automatic update of orthography can be applied to all text fields or to a selected range (e.g. contexts being examples of real texts, should probably be excluded from the spelling update).

Scenario 2: Terminology data collection to be imported

Institution TWO has a terminology database that needs to be enriched with a new domain. There are existing terminological data collections or glossaries on this domain.

Solution	Description	Activities and consequences	Advantages	Disadvantages
1	Add the missing term entries manually.	Modify existing termbase definition, if necessary. Copy and paste all information into terminology database manually. End-users will access a terminology database with more complete information.	Only way of storing and using data that otherwise might get lost.	Notable additional efforts concerning person/months and financial resources. High probability of human input mistakes (typos, omissions, duplicates, etc.) If the termbase definition is not changed to accommodate missing data categories, data might be lost.
2	Automatic import of data.	Modify existing termbase definition, if necessary.	Quick. Can be quite clean,	Can be only relatively clean (missing data, duplicates, interferences

Solution	Description	Activities and consequences	Advantages	Disadvantages
		Prepare data to be imported in format accepted by own terminology database (e.g. XML, TBX, etc.). Perform automatic import of data and subsequent checks for completeness and correctness of import. End-users will access a terminology database with more complete information.	depending on amount and compatibility of data to be imported (e.g. due to completely different data categories and database structure, etc.). Depending on the compatibility of data categories and database structure, limited efforts concerning person/months and financial resources.	with original data, etc.), depending on amount and compatibility of data to be imported (e.g. due to completely different data categories and database structure, etc.). Depending on the availability of data in the necessary format, notable efforts concerning person/months and financial resources. Needs qualified staff able to process data (import/export), bring it into the desired format and modify termbase definition, if necessary.
LISE	Automatic import of data.	Data to be imported and own database prepared and processed by ESTeam, as are checks for	Quick and clean. Limited efforts	Purchase of LISE tool suite.

Solution	Description	Activities and consequences	Advantages	Disadvantages
		completeness and correctness of import (OMEO). End-users will access a terminology database with more complete information. If CLEANUP tool is also used, final terminology database will be cleaner.	concerning person/months and financial resources. Terminologists can view, approve or reject every single addition or merger proposed by the LISE tool OMEO, if desired. Possibility of discussing any addition and merger within the LISE service platform in real time. The CLEANUP tool allows performing further improvements (remove typos, comply to canonical forms, etc.), if desired.	

Solution 1 is an absolutely rare case. It would be chosen only if the desired information were available on paper only or in a format that is not suitable for import into own terminology database. Also, it might be a viable solution for small amounts of data and when there is a high risk of data loss or corruption during automatic import (for any reason).

In solution 2, checks are performed at term level and entry level with the help of the report file usually generated by a terminology management system during import (e.g. a log file).

For solution 2 and 3, it might be necessary to modify the existing termbase definition, e.g. when one of the two terminology data collections contains additional/different data categories. In this case the export file needs to be processed (e.g. substitute data categories) for it to be imported into the modified source terminology database (i.e., where the missing data categories have been added).

If the two terminological data collections have an identical structure and contain the same data categories, no change is necessary, it is possible to proceed directly with importing the new data.

If a new termbase definition is necessary, see scenario 4.

Scenario 3: Terminology data collection to be merged

Institution THREE has two separate source terminological data collections which need to be put together into one target terminological data collection.

Solution	Description	Activities and consequences	Advantages	Disadvantages
1	Add the term entries of one terminology database to the other terminology database manually.	Create new "common" termbase definition, if necessary. Copy and paste all information into terminology database manually. End-users will access only one database instead of two. Doublettes might be created.	Only way of storing and using data that otherwise might get lost.	Notable additional efforts concerning person/months and financial resources. High probability of human input mistakes (typos, omissions, duplicates, etc.) If a new "common" termbase definition is not created to accommodate all data categories, data might be lost. If doublettes are not avoided during input,

Solution	Description	Activities and consequences	Advantages	Disadvantages
				duplicate and partly diverging entries might confuse and frustrate end-users, thus decreasing the trust into the terminology resource.
2	Automatic merger of data.	Create new “common” termbase definition, if necessary. Prepare data to be merged in common format (e.g. XML, TBX, etc.). Perform automatic import of data and subsequent checks for completeness and correctness of import. End-users will access only one database instead of two. Doublettes might be created.	Quick. Can be quite clean, depending on amount and compatibility of data to be merged (e.g. due to completely different data categories and database structure, etc.). Depending on the compatibility of data to be merged, limited efforts concerning person/months and financial resources in preparation and	Can be only relatively clean (missing data, duplicates, interferences between data sets, etc.), depending on amount and compatibility of data to be merged (e.g. due to completely different data categories and database structure, etc.). Depending on the compatibility of data to be merged, notable efforts concerning person/months and financial resources in

Solution	Description	Activities and consequences	Advantages	Disadvantages
			checking phase.	preparation and checking phase. If doublettes are not removed, double and partly diverging entries might confuse and frustrate end-users. Needs qualified staff able to process data (import/export), bring it into the desired format and create new termbase definition, if necessary.
LISE	Automatic merger of data.	Create new “common” termbase definition, if necessary. Data to be imported and own database prepared and processed by ESTeam (OMEO), together with checks for completeness and correctness of import.	Quick and clean. Limited efforts concerning person/months and financial resources. Terminologists can view, approve or reject every	Purchase of LISE tool suite.

Solution	Description	Activities and consequences	Advantages	Disadvantages
		End-users will access only one database instead of two. If CLEANUP tool is also used, final terminology database will be cleaner.	single change or merger proposed by the LISE tool OMEO, if desired. Possibility of discussing any change and merger within the LISE service platform in real time. The CLEANUP tool allows performing further improvements (remove typos, comply to canonical forms, etc.), if desired.	

In solution 2, checks are performed at term level and entry level with the help of the report file usually generated by a terminology management system during import (e.g. a log file).

For solution 1 and 2, it might be necessary to create a new termbase definition which is compatible with the data sets and structures of both terminological data collections, e.g.

a) when the two data sets have different data categories and data structures. Both import files need to be modified so as to obtain uniform data categories; both source data sets (terminological entries) are then imported into a new target terminology database.

b) when one of the two terminology data collections contains obsolete data categories that are not used any more. This becomes the occasion to eliminate the obsolete data categories and clean the pick-lists from items that are not considered useful any more.

c) when updating from an old to a newer terminology management system which has a different way of structuring the data.

If the two terminological data collections have an identical structure and contain the same data categories, no change is necessary, it is possible to proceed directly with importing the new data.

If it is sufficient to modify one of the existing termbase definitions, see scenario 3.

Scenario 4: Terminology databases to be enlarged

Institution FOUR has a terminology database that needs to be enriched with a new language. Terminology work within this institution is translation-oriented. The end-users of the resulting terminological entries work with translation memories and have also already translated a large number of texts into the new language.

Solution	Description	Activities and consequences	Advantages	Disadvantages
1	Slowly add the missing language manually to all entries, as they happen to be worked on (also for other reasons).	Terminology work to search for equivalents, synonyms and any other information needed (e.g. definitions, contexts of use, notes, etc.) in the new language. End-users will access a terminology database with only part of the equivalents available in the new language for a possibly endless amount of time.	Finalised entries will be well-researched and as complete as possible.	Long-term activity that risks being trailed along for months/years. Difficult to determine when data in the new language will be complete and no overview over work already done. Notable additional efforts concerning person/months and financial resources in completing the entries. The termbase will

Solution	Description	Activities and consequences	Advantages	Disadvantages
				contain a mix of entries with and without the new language for a possibly endless amount of time. Missing data may frustrate end-users. Quality of the entries will depend on the quality of terminology work and research.
2	Systematically add the new language to all entries manually.	Terminology work to search for equivalents, synonyms and any other information needed (e.g. definitions, contexts of use, notes, etc.) in the new language. End-users will access a terminology database with only part of the equivalents available in the new language for a limited amount of time.	Finalised entries will be well-researched and as complete as possible. End of activity can be foreseen and planned.	Notable additional efforts concerning person/months and financial resources in completing the entries. The termbase will contain a mix of entries with and without the new language for a limited amount of time.

Solution	Description	Activities and consequences	Advantages	Disadvantages
				Missing data may frustrate end-users. Quality of the entries will depend on the quality of terminology work and research.
3	Automatic extraction from translation memory and update of all entries adding the new language manually.	Automatic extraction of available translations/equivalents in the new language with (commercial) extraction tools. Terminology work to search for equivalents, synonyms and any other information needed (e.g. definitions, contexts of use, notes, etc.) in the new language. End-users will access a terminology database with all equivalents existing in the translation memory in the new language immediately after batch change.	Quick and clean. Limited initial efforts concerning person/months and financial resources (for extraction). Limited additional efforts concerning person/months and financial resources in completing the entries with missing terms.	Reduced additional efforts concerning person/months and financial resources in completing the entries (for terminology work). Entries will only be completed if data is available in translation memory. Quality of the entries will partly depend on the quality of the terminology extraction tool used, of the existing translations and of

Solution	Description	Activities and consequences	Advantages	Disadvantages
				subsequent terminology work and research. Needs qualified staff able to use extraction tools. Purchase or retrieval of extraction tools.
LISE	Batch change and update all entries adding the new language automatically.	Automatic extraction from translation memory of available translations/equivalents in the new language, cross checking of results with already available terminological data; automatic import into the terminology database by ESTeam (FILLUP). Terminology work to search for equivalents, synonyms and any other information needed (e.g. definitions, contexts of use, notes, etc.) in the new language. End-users will access a terminology database with all equivalents	Quick and clean. No additional efforts concerning person/months and financial resources in completing the entries with missing terms. Terminologists can view, approve or reject every single addition proposed by the LISE tool FILLUP, if desired. Possibility of discussing any change and update within the LISE service	Reduced additional efforts concerning person/months and financial resources in completing the entries (for terminology work). Entries will only be completed if data is available in translation memory. Quality of the entries will partly depend on the quality of existing translations and of subsequent terminology

Solution	Description	Activities and consequences	Advantages	Disadvantages
		existing in the translation memory in the new language immediately after batch change. If CLEANUP tool is also used, final terminology database will be cleaner.	platform in real time. In addition to this FILLUP routine, a cleansing step with the CLEANUP tool allows performing further improvements (remove typos, comply to canonical forms, etc.), if desired.	work and research. Purchase of LISE tool suite.

Next to translation memories, also aligned text corpora can be used for bilingual extraction.

FILLUP provides translation suggestions. It locates terms within their context, as part of a translation memory pair. Users can view the pair and extract the correct equivalent from the aligned sentence pair.

7. *Some further, more specific example scenarios*

Scenario 5: Alphabetic ordering of terms

Institution FIVE owns a large terminology database in which the order of the terms within the language sections reflects the chronological sequence of input, not an alphabetic order. As the work is not standardisation-oriented and data needs to be printed in a paper dictionary, there is the need to reorder all terms with their synonyms/variants within each entry alphabetically.

Solution	Description	Activities and consequences	Advantages	Disadvantages
1	Slowly check all entries, as they happen to be worked on (also for other reasons) detecting those that contain synonyms.	Manually reorder synonyms/variants alphabetically. End-users will access a terminology database with entries that have a different ordering of synonyms/variants (alphabetical vs. chronological) for a possibly endless amount of time.	Limited additional efforts concerning person/months and financial resources.	Long-term activity that risks being trailed along for months/years. High probability of human input mistakes. Difficult to determine when activity will be completed.
2	Systematically check all entries, detecting those that contain synonyms.	Systematically detect entries containing synonyms/variants (manual search, specific search tool, filter, etc.) and reorder them	End of activity can be foreseen and planned. Depending on the quantity of the entries	Depending on the quantity of the entries containing synonyms/variants, notable additional

Solution	Description	Activities and consequences	Advantages	Disadvantages
		alphabetically. End-users will access a terminology database with entries that have a different ordering of synonyms/variants (alphabetical vs. chronological) for a limited amount of time.	containing synonyms/variants, limited additional efforts concerning person/months and financial resources.	efforts concerning person/months and financial resources. High probability of human input mistakes.
3	Batch change and update all entries with all synonyms/variants in the new alphabetical order.	The entire terminology database is exported and processed with specific programmes/scripts: entries containing synonyms/variants are detected, reordered alphabetically, and then imported back into the source terminology database. End-users will access a terminology database with all entries containing synonyms/variants in alphabetical order immediately after import.	Quick and clean. Depending on the availability of programmes/scripts, limited efforts concerning person/months and financial resources.	Depending on the availability of programmes/scripts, notable efforts concerning person/months and financial resources. Needs qualified staff able to process data (import/export) and to use specific programme or write necessary script. Purchase/retrieval of programme or writing of

Solution	Description	Activities and consequences	Advantages	Disadvantages
				script.
LISE	Batch change and update all entries with all synonyms/variants in the new alphabetical order.	The entire terminology database is exported and automatically processed by ESTeam: entries containing synonyms/variants are detected, reordered alphabetically, and then imported back into the source terminology database. End-users will access a terminology database with all entries containing synonyms/variants in alphabetical order immediately after import.	Quick and clean. Almost no efforts concerning person/months and financial resources. Terminologists can view, approve or reject every single change and update proposed by the LISE tool CLEANUP, if desired. Possibility of discussing any change and update within the LISE service platform in real time.	Purchase of LISE tool suite.

Dynamic ordering of data values, also upon export/publication, should actually be a feature of the terminology system. However, it is still mostly not the case and the ordering must usually be applied in a separate step.

Scenario 6: Deleting duplicate entries (doublettes)

Institution SIX has realised that its terminology database contains a high number of doublettes, i.e. two or more terminological entries that describe the same concept. This might be due to different possible reasons, e.g. error in import/export routine, import of many different collections of data, several terminologists/institutions working on the same large database, etc. The aim is to remove all doublettes.

Solution	Description	Activities and consequences	Advantages	Disadvantages
1	Do nothing.	End-users will have to disregard or choose from doublettes themselves. The number of doublettes might further increase as time goes by.	No additional efforts concerning person/months and financial resources.	No control over presence of doublettes. Presence of doublettes may confuse end-users, thus decreasing the trust into the terminology resource. Inconsistent language coverage, i.e. the entries each concern (partly) different languages.
2	Slowly delete or merge doublettes manually, as they happen to be found during routine activities/checks.	End-users will have to disregard or choose from doublettes themselves for a possibly endless amount of time.	Limited additional efforts concerning person/months and financial resources.	Long-term activity that risks being trailed along for months/years. Difficult to determine

Solution	Description	Activities and consequences	Advantages	Disadvantages
				when activity will be completed and no overview over work already done. Presence of doublettes may confuse end-users, thus decreasing the trust into the terminology resource.
3	Systematically delete or merge all doublettes	Find and delete or merge all doublettes manually (e.g. with the help of some specific search tool, filter, etc.). End-users will have to disregard or choose from doublettes themselves for a limited amount of time.	Depending on the quantity of doublettes, limited additional efforts concerning person/months and financial resources. End of activity can be foreseen and planned.	Depending on the quantity of doublettes, notable additional efforts concerning person/months and financial resources. Presence of doublettes may confuse end-users, thus decreasing the trust into the terminology resource.
4	Batch change and delete or	The entire terminology database is exported, processed with a	Quick.	Can be only relatively clean, depending on

Solution	Description	Activities and consequences	Advantages	Disadvantages
	merge all doublettes	script/programme to find duplicate entries to be deleted or merged, then imported back. End-users will access a more consistent terminology database immediately after batch change.	Can be quite clean, depending on amount of data to be detected, directly deleted or merged with other entries. Depending on the availability of scripts/programmes as well as on the amount of data to be deleted or merged with other entries, limited efforts in post-processing.	amount of data to be detected, directly deleted or merged with other entries. Depending on the availability of scripts/programmes as well as on the amount of data to be deleted or merged with other entries, notable efforts in post-processing. Difficult to keep track of what should be deleted, kept, merged. Needs qualified staff able to process data (import/export) and to write script or run programme. Purchase/retrieval of programme or writing of

Solution	Description	Activities and consequences	Advantages	Disadvantages
				script.
LISE	Batch change and delete or merge all doublettes	The entire terminology database is exported, automatically processed by ESTeam, changes are viewed and approved/rejected by own terminologists, then applied by ESTeam (OMEO). End-users will access a more consistent terminology database immediately after batch change. If CLEANUP tool is also used, final terminology database will be cleaner.	Quick and clean. Limited efforts concerning person/months and financial resources. Terminologists can view, approve or reject every single deletion or merger proposed by the LISE tool OMEO, if desired. Possibility of discussing any deletion and merger within the LISE service platform in real time. The CLEANUP tool allows performing further improvements (remove typos, comply to canonical forms, etc.), if desired.	Purchase of LISE tool suite.

In solution 2 and 3, it might be necessary to collect the feedback of the departments/institutions that originally created the terminological entry to obtain their approval for deletion or merger with another entry. This would have to be done in physical meetings, via email or other communication channels, in a step that is separated from the terminological process.

In solution LISE, the discussion and approval/rejection can take place within the LISE platform.

For solution 4, note that up until recently most commercial products did not allow automatic detection/filtering of doublettes.

Scenario 7: Controlling synonyms and variants

Institution SEVEN wishes to standardise its multilingual language use or introduce a controlled use of terminology (in one or more languages) and needs to find the preferred term for all entries and languages. The aim is to mark which synonyms/variants are to be used and which ones avoided as well as deleting some superfluous synonyms/variants (if considered necessary)..

Solution	Description	Activities and consequences	Advantages	Disadvantages
1	Slowly detect all entries that contain synonyms/variants, as they happen to be worked on (also for other reasons).	Terminologists will retrieve all entries containing synonyms/variants in the source and target language(s) and delete inacceptable synonyms/variants or mark the preferred terms as such. End-users will access a terminology database with entries that have a different treatment of synonyms/variants for a possibly endless amount of time.	Limited additional efforts concerning person/months and financial resources.	Long-term activity that risks being trailed along for months/years. Difficult to determine when activity will be completed and no overview over work already done.
2	Systematically retrieve all entries that contain synonyms/variants.	Terminologists will systematically retrieve (either manually or with some specific search tool, filter, etc.) all entries containing synonyms/variants in the source	End of activity can be foreseen and planned.	Notable additional efforts concerning person/months and financial resources.

Solution	Description	Activities and consequences	Advantages	Disadvantages
		and target language(s) and delete inacceptable synonyms/variants or mark the preferred terms as such. End-users will access a terminology database with entries that have a different treatment of synonyms/variants for a limited amount of time.		
3	Automatically retrieve all entries that contain synonyms/variants.	The entire terminology database is exported and processed with a script/programme to retrieve all entries that contain synonyms/variants. Terminologists delete inacceptable synonyms/variants or mark the preferred terms as such. End-users will access a terminology database with entries that have a different treatment of synonyms/variants for a limited amount of time.	Quick and clean. Depending on the availability of scripts/programmes as well as on the complexity of data, limited efforts concerning person/months and financial resources.	Depending on the availability of scripts/programmes as well as on the complexity of data, notable efforts concerning person/months and financial resources. Notable efforts concerning person/months and financial resources in post-processing (deleting synonyms and marking

Solution	Description	Activities and consequences	Advantages	Disadvantages
				preferred terms). Needs qualified staff able to process data (export) and to write script or run programme. Purchase/retrieval of programme or writing of script.
LISE	Systematically retrieve all entries that contain synonyms/variants automatically.	The entire terminology database is exported and automatically processed by ESTeam: entries containing synonyms/variants in the source and target language are detected to allow processing and then re-imported into source terminology database. Terminologists delete inacceptable synonyms/variants or mark the preferred terms as such within the LISE tool, following or discarding the suggestions provided by the	Quick and clean. Limited efforts concerning person/months and financial resources. Terminologists can view, approve or reject every single change and deletion proposed by the LISE tool OMEO, if desired. Possibility of discussing any change or deletion	Purchase of LISE tool suite.

Solution	Description	Activities and consequences	Advantages	Disadvantages
		tool. End-users will access a terminology database with all entries containing only the desired synonyms/variants (or with the preferred terms marked as such) immediately after import. If CLEANUP tool is also used, final terminology database will be cleaner.	within the LISE service platform in real time. The CLEANUP tool allows performing further checks (typos, canonical forms, etc.) and corrections, if desired.	

For scenario 1, 2 and 3 note that the process of decision-making on the preferred terms can be time-consuming, since it usually requires meetings or lengthy correspondence via email. In scenario 4 this can be discussed directly within the platform.

Scenario 8: Monitoring language coverage

Institution EIGHT working with many languages and owns a large collection of terminology data wishes to know which entries lack one or more languages (e.g. to subsequently proceed with scenario 4).

Solution	Description	Activities and consequences	Advantages	Disadvantages
1	Slowly detect the missing language(s) manually in all entries, as they happen to be worked on (also for other reasons).	Manually mark, tag or list all entries where one or more languages are missing.	Limited additional efforts concerning person/months and financial resources.	Long-term activity that risks being trailed along for months/years. Difficult to determine when activity will be completed and no overview over work already done. Difficult to plan future integrations and completion.
2	Systematically detect the missing language(s) in all entries.	Systematically mark or list all entries where one or more languages are missing. Future integrations and completion	End of activity can be foreseen and planned.	Notable additional efforts concerning person/months and financial resources.

Solution	Description	Activities and consequences	Advantages	Disadvantages
		can be planned.		
3	Automatically detect the missing language(s) in all entries.	The entire terminology database is exported and processed with a script/programme to detect missing languages in entries. Future integrations and completion can be planned.	Quick. Depending on the availability of scripts/programmes as well as on the complexity of data, limited efforts concerning person/months and financial resources.	Depending on the availability of scripts/programmes as well as on the complexity of data, notable efforts concerning person/months and financial resources. Needs qualified staff able to process data (export) and to write script or run programme. Purchase/retrieval of programme or writing of script.
LISE	Automatically detect the missing language(s) in all entries.	The entire terminology database is exported and processed by ESTeam to detect missing languages in entries. Future integrations and completion	Quick. Limited efforts concerning person/months and financial resources.	Purchase of LISE tool suite.

Solution	Description	Activities and consequences	Advantages	Disadvantages
		can be planned.	Possibility of discussing any result and future step within the LISE service platform in real time.	

For solution 2, note that some commercial terminology management systems allow searching a specific language, but not filtering data according to the language. In this way it becomes quite impossible to detect which entries do not contain a specific language.

8. Attachment A – Interview protocol

GENERAL ASPECTS

For how many years have you been dealing with terminology?
Who does terminology work (translators, professional terminologists, others, etc.)?
Are there native speakers for all languages managed?
What kind of education/training do these people have (degree in translation/terminology, specific domain…)?
How many staff members (full time/part time) are there in the terminology service/do terminology work?
When is terminology work necessary?
Why does the institution/company, etc. manage (or not manage) its terminology?
Who are the target users (e.g. internal personnel/public/customers; language experts/domain experts, etc.)?

METHODOLOGY

Is the approach descriptive or prescriptive?
Is terminology work done systematically or ad-hoc?
When is it started (a posteriori/a priori, before product/service development or only before translation)?
What types of domains are considered? How about social security terminology?
How many (and which) languages are considered?
How many (and which) legal systems are considered?
How do you work when dealing with (many) different legal systems (legal comparison, indication of geographical usage in the case of legal systems using the same language, etc.)?
Who solves monolingual/multilingual terminology issues (translators, terminologists, domain experts, etc.)?
Who creates/discusses/validates neologisms?
Which steps are done automatically/manually (e.g. terminology extraction)?
Which material/documents is the work based on?
Are internal/freely available electronic corpora used?

TERMINOLOGY MANAGEMENT

Is the work coordinated centrally in one service or are several departments collecting their own terminology collections?
Who is in charge of terminology management?
Who revises it (senior terminologists, domain experts, etc.)?
How are entries corrected/updated (manually/automatically, selectively/batch changes)?

Who is the work in the legal/administrative area done for (executive/legislature/judiciary/administration)?
Which roles are clearly defined? Are there any?
How are these roles defined (e.g. Are terminologists there to feed a database? To create, organise and centralise the terminology?)?
Is quality checked? If so, how and by whom?
Are there specific (written) internal/public guidelines for some or all of the aspects related to terminology work?
Is there one (or more) standard workflow(s) for terminology production?
What does it look like? Which tools are used and how are they employed?
What kinds of problems do you face with this/these type(s) of workflow, if any?
Are the results published internally/to the public?
Who has access to the terminology (only terminology/translation department or the entire institution/the public)?
How are new languages added (manually/automatically, selectively/batch changes)?
Harmonising or merging databases: how do you proceed (manually, semi-automatically)? What are the workflow steps and what problems do/did you face, if any?

WORKING WITH TMS

Is there an internal/online termbase?
What new data/domains/languages are planned for the future?
What data is kept, merged, deleted, updated in a termbase?
Are data only managed or also enriched and added?
Where are data stored? Locally, centrally, online? In what system?
What does the data model look like?
What does the entry structure look like?
What other data are available (e.g. corpora, internal documentation, etc.)?
Which tools would you use (besides the ones maybe already used) for data clean-up, merging, fill-up?

TERMINOLOGY PLANNING

Is the work translation-oriented? Standardisation-oriented? Multifunctional?
Is the collected terminology standardised by some responsible (body)?
Are there specific guidelines for terminology standardisation?
How is the standardisation process organised (commission, experts, meetings between terminologists and experts, online communication/forum, etc.)?

9. Attachment B – Online questionnaire

Question 1: For how many years has your institution been doing terminology work?
- 0-5 years
- 6-10 years
- more than 10 years

Question 2: Who takes care of terminology work? (check all that apply)
- translators/interpreters
- terminologists
- domain experts
- other (please specify)

Question 3: What kind of education/training do these people have? (check all that apply)
- university degree in translation
- university degree in terminology
- professional training in translation
- professional training in terminology
- university degree in a specific domain (e.g. law, medicine, etc.)
- other (please specify)

Question 4: How many full time staff members are there in the terminology service/do terminology work? (please include also part-time positions, e.g. if one person works full time, one person part time at 60% and one person part-time at 25%, answer 1.85)

Question 5: Who are the target users of your terminology? (check all that apply)
- internal staff
- customers
- general public
- translators, interpreters, other language experts
- domain experts
- other public/private institutions
- other (please specify)

Question 6: What kind of approach to terminology work prevails at your institution?
- descriptive approach
- prescriptive approach

Question 7: How is the terminology work at your institution oriented?
- text production oriented
- translation oriented
- standardisation oriented
- multifunctional

Question 8: How do you elaborate terminology as a rule?
- systematically (i.e. domain by domain)
- ad-hoc (i.e. when there are specific needs)

Question 9: When is terminology work done as a rule?
- a priori (i.e. before document production or service/product development)

a posteriori (i.e. after document production or service/product development)
during the entire drafting/translation/production process

Question 10: What are the domains your institution mainly works on? (check all that apply)

law & administration
science & technology
economy & finance
medicine
other (please specify)

Question 11: For whom is the terminology work in the legal and administrative area mainly done? (check all that apply)

executive
legislature
judiciary
administration
other (please specify)

Question 12: How many legal systems are considered?

Question 13: Which legal systems are considered?

Question 14: Do you work or have you worked on social security terminology?

yes
no

Question 15: Are you planning to add new domains in the future?

yes
no
maybe

Question 16: Which new domains are you planning to add in the future? Question 17: How many languages do you work with?

Question 18: Which languages do you work with? (check all that apply)

Basque	Dutch	Galician	Latvian	Slovak
Bulgarian	English	German	Lithuanian	Slovenian
Catalan	Estonian	Greek	Maltese	Spanish
Croatian	Finnish	Hungarian	Polish	Swedish
Czech	French	Irish	Portuguese	Welsh
Danish	Gaelic	Italian	Romanian	

Question 19: Are you planning to add new languages in the future?

yes
no
maybe

Question 20: Which new languages are you planning to add in the future? (check all that apply)

See question 18

Question 21: Do you employ native-speaker terminologists for all the languages you work with?

yes

no

Question 22: Who solves monolingual terminology issues as a rule? (check all that apply)

translators/interpreters

terminologists

domain experts

others (please specify)

Question 23: Who solves multilingual terminology issues as a rule? (check all that apply)

See question 22

Question 24: Who creates neologisms as a rule? (check all that apply)

See question 22

Question 25: Who validates neologisms as a rule? (check all that apply)

See question 22

Question 26: On which material/documents is your terminology work based? (check all that apply)

internal documentation

legislation

manuals, handbooks, etc.

technical reports & documentation

websites (containing other than legislation, manuals, etc.)

other (please specify)

Question 27: Do you use corpora, i.e. large collections of authentic texts gathered in electronic form according to a specific set of criteria, for your terminology work?

1 (never)

2

3

4

5 (always)

Question 28: The corpora you use are (check all that apply):

freely available

commercial

internal

other (please specify)

Question 29: The corpora you use are (check all that apply):

monolingual

bilingual/multilingual

Question 30: Are the bilingual/multilingual corpora you use aligned (parallel)?

yes

no

Question 31: How do you do terminology extraction as a rule?

manually

semi-automatically

we never do terminology extraction

Question 32: How is terminology work distributed in your institution? (check all that apply)
- centrally in one service
- several departments collect their own terminology
- other (please specify)

Question 33: Where are the terminological data stored? (check all that apply)
- locally
- centrally
- other (please specify)

Question 34: Who is in charge of terminology management? (check all that apply)
- no one
- translator/interpreter
- terminologist
- senior terminologist
- terminology manager
- domain expert
- other (please specify)

Question 35: Which roles are clearly defined in your institution? (check all that apply)
- no roles defined
- trainee
- translator/interpreter
- terminologist
- senior terminologist
- terminology manager
- domain expert
- other (please specify)

Question 36: Who revises terminology work? (check all that apply)
- translator/interpreter
- terminologist
- senior terminologist
- domain expert
- other (please specify)

Question 37: Who does the quality check? (check all that apply)
- no one
- translator/interpreter
- terminologist
- senior terminologist
- terminology manager
- domain expert

Question 38: How are terminological entries corrected/updated? (check all that apply)
- manually/selectively
- automatically/batch changes

Question 39: Who chooses the termbase definitions (i.e. data categories to be included, structure of the database, etc.)? (check all that apply)
- translator/interpreter

terminologist
terminology manager
computational linguist/computer expert
no experience
other (please specify)

Question 40: Who manages the import/export of terminology data? (check all that apply)
translator/interpreter
terminologist
terminology manager
computational linguist/computer expert
other (please specify)

Question 41: Do you have one (or more) standard workflow(s) for terminology production?
yes
no

Question 42: Are there specific guidelines for some or all of the aspects related to terminology work?
yes
no

Question 43: Are these guidelines accessible? (check all that apply)
for internal staff only
freely available
only partly available to the public
other (please specify)

Question 44: Who has access to the results of your terminology work?
all the staff and some selected institutions/people
all the staff and the general public
other (please specify)

Question 45: Does the general public see the same (amount of) information as your internal staff?
yes
no

Question 46: Do you use any specific tools for your language-related work (e.g. computer-assisted translation tools, terminology extraction tools, terminology management systems, etc.)?
yes
no

Question 47: What kind of tools do you use? (check all that apply)
translation memory (TM)
terminology management systems (TMS)
terminology extraction software
controlled language checker (CLC)
authoring memory tool
translation workflow tool
other (please specify)

Question 48: The tools you use are mainly (check all that apply):
commercial
developed in-house

commercial but adapted for specific needs
freely available
other (please specify)

Question 49: How are/were new languages added to the terminology collection? (check all that apply)
no experience
manually/selectively
automatically/batch changes

Question 50: How do/did you proceed when merging databases? (check all that apply)
no experience
manually
semi-automatically
automatically

Question 51: Do you use any standard terminology exchange format (e.g. TBX)?
yes
no

Question 52: Do you standardise terminology?
yes
no

Question 53: Are there specific guidelines for terminology standardisation?
yes
no

Question 54: Who decides/decided the criteria for terminology standardisation? (check all that apply)
there are no specific criteria defined
terminologists
standardisers

Question 55: Who standardises terminology? (check all that apply)
standardisation commission
domain experts
terminologists

Question 56: How is the standardisation process organised? (check all that apply)
physical meetings
online meetings (e.g. videoconferencing)
online discussion forum
other (please specify)

Question 57: Would you like to receive the results of the questionnaire and/or of the LISE project? If so, please indicate your affiliation and email address below:

Question 58: Do you have any comments?