Quality assurance in multilingual legal terminological databases
Elena Chiocchetti, Eurac Research, Tanja Wissik, Austrian Academy of Sciences, Vesna Lušicky, University of Vienna, and Michael Wetzel, Coreon GmbH

ABSTRACT
In addition to several other factors, terminology and terminology management are perceived to be important elements of the quality assurance (QA) system of document production, translation, and multilingual lawmaking. This implies that quality in translation, including legal translation, partly depends on the quality of terminology found in terminological databases (TDBs). In order to use terminology as a QA tool, the quality of TDBs has to be ensured. In this paper, we discuss the relevance of legal TDBs for translation quality and propose a QA framework for multilingual legal TDBs based on a comprehensive approach, which includes QA at workflow level, at product level, and at staff level. For each of these three aspects we address the main features to be implemented — and how they should be implemented — to successfully achieve and maintain high quality of multilingual legal TDBs. Our comprehensive approach to QA therefore considers persons, processes, products and services as well as dedicated tools.

KEYWORDS
Quality assurance, terminology, legal terminological databases, term bases, legal translation, quality standards, quality assurance framework.

1 Introduction
Aspects and challenges in translation quality have received substantial attention in the translation profession and academic environment (cf. inter alia House 1977, 1997, 2015; Schäffner 1997; Brunette 2000; Lauscher 2000; Schippel 2006; Colina 2008; Kelly and DePalma 2008; Schmitt et al. 2009; Williams 2009; O’Brian 2012). Three terms have been often used interchangeably to refer to quality related activities: quality assurance (QA), quality assessment, and quality control (for details on each term cf. section 4). In addition to several other factors, terminology and terminology management are perceived to be important elements of the QA system of document production (Valentini 2016), translation (Popiolek 2015), and multilingual lawmaking (Strandvik 2015). This implies that quality in translation, including legal translation, partly depends on the quality of terminology and terminology management. Using terminological databases during a translation assignment “ensure[s] that uniform, consistent terminology is used throughout a translation or by a project team and can make a significant contribution to the quality of a translation” (Risku 2006). In this paper, we therefore focus on QA in multilingual legal terminological databases. We will discuss the different aspects of QA at the level of workflow, product and staff.

2 The relevance of terminological databases for translation quality
Terminological databases (TDBs) provide “a structured repository of linguistic data, enriched with metadata” (Steurs et al. 2015: 224). Their
structure follows predefined classifications (e.g. using an established subdivision of the legal domain into subdomains, such as public law, criminal law, labour law, etc.) and relies on concept based analyses of the terminology pertaining to one or more specialised domains (cf. Steurs et al. 2015: 224).

TDBs offer ready-made terminological solutions to the translators who consult them: they collect established equivalents (Molina and Hurtado Albir 2002: 510, Biel 2008: 26), thus ensuring smooth understanding and sparing the time otherwise necessary for terminological research. This is especially important for legal terminology as it relieves translators not only from the task of searching equivalents, which might per se be quite demanding, but also from actual comparative work (i.e. micro-comparison at concept level), which requires in-depth legal and subject matter knowledge.

TDBs provide translators with more information than specialised dictionaries. Definitions and comparative notes in TDBs for instance are useful to fill potential knowledge gaps of translators in both source and target language (cf. Magris 2004). TDBs also place terms within the conceptual system of the specialised domain treated, for example, when definitions specify the genus proximum or when there are detailed references to hyperonym, co-hyponyms, hyponyms and otherwise related terms. When the equivalents in TDBs are more context based rather than concept based, the information given allows the translator to determine whether the proposed equivalent is suitable for the text, such as the degree of equivalence or the specific context of use. Also in case of homonyms or polysemous terms, the user finds the necessary information to choose the right term and the most adequate translation.

Translators not only consult available TDBs but also produce and feed them during their work in order to avoid repeating time-consuming search activities which would curb productivity (cf. Bowker 2015: 306). In fact, the time dedicated to terminological research can make up a substantial amount of time needed for a translation assignment. Recent estimates vary form 20-25 % to 40-60 % according to the degree of experience of the translator (cf. Désilets et al. 2009, Gomez Palou Allard 2012, Champagne 2004 in Bowker 2015: 311). Efficient terminology management therefore helps translators improve the linguistic quality of their texts, reduces costs and shortens the time needed to complete a translation assignment (cf. Bowker 2015: 305).

TDBs are handier and often more effective in addressing translators’ problems than traditional dictionaries (cf. Magris 2004: 55-56). Being electronic tools, they are quick and easy to consult and allow extended searches (e.g. by clicking on related terms to view the respective entries). Many TDBs can be directly accessed from computer assisted translation tools (CAT tools). It is also possible to share them, i.e. send them to
translators for reference during specific assignments or exchange them between colleagues.

TDBs are a quality assurance tool in translation. They ensure that terminology is used correctly and consistently within and across texts, both in the source language (also for controlled language use) and in the target language. Quality assurance is actually considered one of the main goals of terminology work (cf. Schmitz and Straub 2010: 49ff).

In the translation QA process, terminology is relevant before, during and after translation (i.e. during revision), as it is an element which must be checked for compliance with set standards (cf. Popiołek 2015: 345). This is usually done with the help of a TDB, a CAT tool (its main component being the translation memory), and a quality assurance module. The TDB should ideally be integrated in the CAT tool or interact well with the tool, so that searches can be automated and important terminological information becomes immediately visible while the translator is working on an assignment with the help of a translation memory. The QA module then serves to check whether the desired terminology was used (cf. Popiołek 2015: 347).

3 Characteristics of legal terminology and databases of legal terminology

When dealing with TDBs that collect legal terminology it is important to consider some essential characteristics of legal terminology, which make it particularly difficult for translators to first understand and then translate legal texts. Legal terms have a high degree of abstraction (e.g. ‘legal personality’). In fact, there are very few concrete legal objects, so that legal systems with their sets of concepts, relations and rules are actually created and modified by the legal language rather than being merely described by it (cf. Fioritto 2007: 408).

Legal terminology may also be characterised by some degree of vagueness. For example, for concepts that need to be interpreted and applied to specific situations (e.g. ‘indecent circumstances’). This is often functional and useful (cf. Prandi 2010) as it allows laws to be applicable in a wide range of situations and adapt to the evolution of society without being constantly revised and rewritten.

An in-depth analysis of databases collecting legal terminology will reveal a heterogeneous, albeit always legally relevant set of terms. Next to what we might call legal terminology proper, i.e. terms that are typical and exclusive to the legal domain (e.g. ‘subornation of perjury’), there will be special language terms from other domains that are relevant for lawmaking on specific subjects (e.g. ‘surrogacy’, ‘noise pollution’). These terms have a double specialisation as they still belong to the original domain but have also become legal terms (cf. Soffritti 2002: 60). We will also find common
language words redefined to assume a legal meaning (e.g. ‘trust’) (cf. Rega 2002: 54), legal neologisms (e.g. ‘telework’, ‘job-on-call’) as well as loan words (e.g. ‘acquis communautaire’, ‘referendum’). An important role is also played by collocations (e.g. Noun + Verb combinations like ‘to terminate a contract’, ‘to fulfil an obligation’) and formulaic expressions (e.g. ‘Now therefore, the Parties have agreed as follows’). The latter can even be longer phrases, often called boilerplates (e.g. ‘All claims and disputes arising under or relating to this Agreement are to be settled by binding arbitration in the state of X or another location mutually agreeable to the Parties.’), which basically recur in almost identical form in specific legal text types, for example in contracts and judgements (cf. Kjaer 2007, Biel 2014). Such textual building blocks may be stored in contexts within legal TDBs.

Synonymy and term variation are frequent phenomena in the legal domain (cf. Lavagnino 2010) that need to be accounted for in dedicated TDBs (e.g. ‘causal link’, ‘causal nexus’ or ‘causal connection’; ‘plaintiff’ – now ‘claimant’). The same holds true for homonyms and polysemous terms (e.g. ‘party’ meaning ‘political party’ or a ‘party’ in court such as the claimant).

The most important characteristic of legal terminology is its indissoluble bond with the legal system it belongs to, which de Groot (1999:12ff) terms Systemgebundenheit. Consequently, full equivalents are rare across different legal systems (cf. Sandrini 1996:138, Šarčević 1997:232, 236ff) (e.g. the Italian certificazione antimafia, which companies need when participating in public tenders to prove that they are not involved in criminal activities, has no equivalent in other systems). Differences at conceptual and terminological level exist even when comparing legal systems that use the same official language (cf. Gambaro and Sacco 1996: 9) (e.g. collective agreements in Germany are called Tarifvertrag, in Austria Kollektivvertrag). The search for equivalents in other legal systems therefore implies not only terminological comparison but also legal comparison methods (micro-comparison) (cf. Chiocchetti et al. 2013a: 12ff; 2013b: 11ff).

All features mentioned above are relevant for the structure and content of a TDB of legal terminology. Data presentation in a TDB not only depends on content, but also on many other aspects (Steurs et al. 2015: 227). The purpose and target audience of the terminology collection are of paramount importance in this respect. For example, translation-oriented TDBs might contain less detailed conceptual information but more collocations, formulaic expressions and usage notes than a standardisation-oriented TDB. Similarly, a collection aimed mainly at supporting translators will present different types of information than a TDB aimed primarily at domain experts. Definitions in a TDB with legal content for instance may need to be more exhaustive for non-experts while giving more legal cross references for domain experts. TDBs actually often address multiple targets (e.g. domain experts, drafters, translators, students) with the ensuing difficulty of presenting data in a complete yet useful format for a diversified audience.
The number of legal systems considered in the TDB is also relevant. A multilingual database referring to only one legal system — and therefore presenting the same concepts in different languages (as might happen in multilingual countries like Switzerland) — will face different issues with respect to a TDB containing terms of two or more legal systems that need to be compared (e.g. a database envisaged to support transnational cooperation). While the first TDB will deal primarily with different conceptualisation, designation and term formation strategies, diverging associations, etc., across languages, the second one will be confronted with the substantial differences between legal systems in addition to all the aforementioned aspects. The latter TDB should clearly identify the legal system every single term belongs to. This is especially important when the database contains terms from different legal systems using the same language (e.g. German in Germany and Austria). Another issue that TDBs with legal content need to address are the strategies adopted to illustrate the degree of equivalence between legal terms and the concepts they designate (e.g. explanations in fully fledged comparative notes, standard codes or symbols).

Term retrieval may also pose problems, since domain attribution is to a certain extent arbitrary. Even though the classification of terms into domains and subdomains might follow standard subdivisions (e.g. the conventional categorisation of legal subjects), these are not necessarily shared by all legal systems and many terms might be reasonably attributed to multiple related subfields. In addition, users, especially translators, might not necessarily be familiar with conventional subdivisions. Finally, there is the challenge of dealing with homonyms and polysemous terms in a clear and consistent way so that users may understand immediately which terminological entry to select in the hit list of search results.

All issues described in this section pose challenges to multilingual databases of legal terminology. The quality of a TDB not only depends on the quality of its content but also on the fact that all above mentioned issues are considered and solved consistently and clearly. In the next sections, we will therefore explain how to improve the quality of TDBs in order to make them easier to use, trustworthy and valuable for legal translators.

4 Quality assurance in terminology work

As outlined in the introduction, quality aspects and quality models in translation are well covered topics in literature. In addition, several standards and guidelines have been developed to adapt the models for the industry’s needs (cf. section 4.1 for details). Despite the standards and extensive literature on the topic, terminology used in reference to aspects of quality management remains ambiguous. The most widely used terms are quality assurance (QA), quality assessment (especially in regard to translation quality assessment (TQA)), and quality control (QC) — all being
key elements in quality management systems. The first term, QA, underlines a proactive approach and refers to “a model approach that ensures good results if the right combination of human and technical resources is used in a sequence of steps and tasks that constitute a process within a system” (Popiołek 2015: 341). Quality assessment, sometimes quality evaluation, is the measurement of the extent to which a product or service complies with quality specifications. Quality control encompasses checking whether the products or services meet stated quality specifications (Lommel (ed.) 2015). Quality assessment and quality control thus provide only the verification of the compliance with the planning and preventive measures set out in QA.

In the absence of an absolute definition of quality, the concept of quality in the translation industry has been developed to be client and application driven (cf. Budin 2007). To address the need for objective, comparable and measureable parameters for QA, TQA and QC while fulfilling diverse client requirements, the translation industry often focuses on the three P’s of quality assessment (cf. Stejskal 2006, 2009): product, process and provider. As discussed above, terminology and TDBs influence the quality of translation and are often an important element in translation QA. To ensure that the terminology used in translation is of adequate and accepted quality, we propose adopting a comprehensive approach for QA in multilingual legal TDBs, in which processes, products/services and people are the three key elements of QA.

This also implies that the quality of terminology work itself has to be proactively managed, assessed, monitored and controlled. However, QA in terminology work has only recently become a more relevant topic. In the past, it received little attention or was only addressed indirectly (e.g. Wright 2001: 488ff, Fähnrich 2005, Van den Bogaert 2008, Cerrella Bauer 2009). Recently, Kockaert and Steurs (2015) dedicated a whole section in the Handbook of Terminology to QA. Only a few authors have focused more in depth on quality related activities in terminology management or QA of terminological resources (e.g. de Coronado et al. 2009, Kudashev 2013, Carlson et al. 2014, Valentini 2016). Carlson et al. (2014: 14) state that “an important part of any framework for creating and managing terminology is quality assurance infrastructure”. Kudashev (2013: 15) defines such a QA infrastructure as “facilities designed to assure the specified quality level of terminological data and terminology management operation.” Kudashev’s QA infrastructure mostly concentrates on the product level, i.e. the TDB:

[q]uality assurance infrastructure in a term bank mostly consists of various types of metadata. Depending on its function and the object to which it relates, metadata can be divided into structural and descriptive metadata. Another important element of quality assurance infrastructure is methodological data (Kudashev 2013: 15).
Workflows and user roles have so far been a marginal topic from the QA perspective, although Cerrella Bauer (2009, 2015) discusses them from the industry perspective.

In the following, we will discuss international standards and their relation to QA in terminology work.

4.1 Quality assurance and standards

Providers of multilingual legal TDBs are usually an entity in a regional, national, international or supranational organisation with a structured translation service. Many of these organisations have adopted a quality management system, for example ISO 9001:2015, Quality management systems – Requirements from the range of ISO 9000 standards, Total Quality Management (TQM), Six Sigma or a similar quality management system. Quality in the scope of the ISO 9000 family refers to the “degree to which a set of inherent characteristics [...] of an object [...] fulfils requirements” (ISO 9000:2015, section 3.6.2). In short, it argues that quality cannot be achieved in a vacuum, but that it is ultimately determined by the users and applications of the product or service. The identification of relevant stakeholders, a needs analysis and the identification of requirements are well-established steps in the terminology project cycle of corporate terminology (Cerrella Bauer 2009, 2015). They have however proven sometimes difficult to implement in settings with large stakeholder groups, or in projects setting up terminological databases available to the general public, which is often the case in the legal domain (cf. Lušicky and Wissik 2012). In general, adopting the ISO 9001 principle is possible in four steps: explicitly say what you do, do what you say, prove what you do, and document what you do. Adequate terminology management tools and quality management can support these steps, especially in highly regulated environments, such as the legal domain (Lušicky and Wissik 2015: 65).

Standards have become a prerequisite for ensuring quality through systematic quality management. Quality and quality management in specialised communication have been implicitly and explicitly addressed by a number of international standards that are primarily developed in the framework of standardising bodies, mainly by international standards organisations, among others by the International Organization for Standardization (ISO) and the European Committee for Standardization (CEN) (cf. Budin 2007).

address various aspects of products and processes in terminology work. In addition, possible data categories for terminological entries are specified in ISO 12620:2009, *Terminology and other language and content resources – Specification of data categories and management of a Data Category Registry for language resources*, and more specifically in ISO 12612:2002, *Translation-oriented terminography*, which includes a selection of the data categories considered to be of particular relevance to the translation process. In the light of the many differences among terminological databases and tools and hence interoperability problems, ISO 30042:2008, *Systems to manage terminology, knowledge and content – TermBase eXchange (TBX)*, provides an instrument to address these challenges (cf. Melby 2012).

Terminology standards are also referenced in several standards that support what might be referred to as the QA system for a translation process, among others: ISO 17100:2015, GB/T 19682-2005, SAE J2450 (2005).^2^ In the following we will describe the three aspects of the terminology QA in more detail.

### 4.2 Quality assurance at workflow level

There are many models for terminology workflows and processes, as well as many ways to systemise the different steps and tasks involved, because the workflows depend on the type of terminology work performed and on the institutional environment.

Based on the definition of translation workflow in ISO 17100: 2015, the terminology workflow can be defined in analogy as “the process or parts thereof involved in elaborating and publishing terminological data”. Most terminology workflows involve QA steps in some way (e.g. *inter alia* COTSOES 2002; Deutscher Terminologie-Tag e.V. 2010; Schmitz and Straub 2010; Chiocchetti *et al.* 2013a, 2013b; Lušicky and Wissik 2015; Popiołek 2015).^3^ However, there are also descriptions of terminology workflows where QA is not explicitly mentioned (e.g. Arntz *et al.* 2014). The Best Practice Guide of the Deutscher Terminologie-Tag e.V. (2010: M5-7/8) describes the following steps for the elaboration of terminological data, especially in the context of industry: production of terminological data (*Produktion*), preparation and publication of terminological data (*Bereitstellung*), use of terminological data (*Nutzung*) and QA (*Qualitätssicherung*). De Coronado *et al.* (2009) report on the QA steps during the different phases of editing and publication of a biomedical terminological resource. Chiocchetti *et al.* (2013a: 14ff) describe a prototypical workflow for legal terminology work. This workflow involves the following steps, which are not always all performed: needs analysis, defining priorities, documentation, term extraction, term selection, elaboration of terminological entries (with contrastive analysis and micro
comparison), revision and QA, and dissemination. The final product resulting from the workflow may serve translators and other users, but terminology use is not necessarily a step within the terminology workflow sensu stricto.

Even though workflow charts may give the impression that QA is only one step in the workflow, it has to be pointed out that QA has to be involved in all the different workflow steps on a routine basis, and not only at the end of the workflow. De Coronado et al. (2009) state that “[a] variety of different QA steps and processes are conducted both routinely during each production cycle, and on a periodic basis as ongoing QA.”

In section 4.1 we have seen that several standards address QA, but there are no specific process standards for terminology work and neither are there for legal terminology work.

Since terminology can be seen as a translation issue (cf. Warburton 2006), terminology work also falls under the scope of the wider framework of translation services, and may thus be subject to provisions required for translation services as defined in ISO 17100: 2015 (and before in EN 15038: 2006). This standard also mentions terminology management in the list of value added services (ISO 17100: 2015, Annex F). Furthermore, Popiołek (2015) illustrates how terminology management is embedded in the translation QA process. Strandvik (2015) shows that ISO 17100:2015 has a wider scope by applying it to multilingual lawmaking. The ISO 17100 divides the workflow in a pre-production process and activities, production process and post-production processes, and sets conditions for achieving quality by following the defined process steps and activities. Therefore, the accountability of each process in terminology work is of uttermost importance and is ensured by relevant documentation (ISO 17100:2015, Lušicky and Wissik 2015):

- for handling and analysing enquiries, and for determining project feasibility,
- for determining whether all human and technical resources are available,
- for handling project-related information, resources, and documentation,
- for all terminology-related activities (e.g. documentation of source quotations),
- for the collection, assessment, traceability, and follow-up of user satisfaction,
- by an agreement with the client or by a similar agreement.

Furthermore, all activities related to terminology should be subject to the overall documented quality management system that is in place in the organisation.
The above mentioned documentation exists for example as guidelines. De Coronado et al. (2009: 532) state that a guideline or manual "provides the first line of QA," and also stress that “[i]t is also important to review the guidelines periodically, to be sure that they are applied as intended and that they are still valid.”

QA at process level can also be achieved by workflow management tools, which streamline the terminology processes and keep track of ongoing or finished tasks, thus also documenting the process. Even though there are many workflow management tools on the market, specific tools for terminology workflows are quite rare. One tool for example is quickTerm by Kaleidoscope (Kaleidoscope 2015), a workflow tool for requesting, voting on, approving, and translating terms. The tool also includes an email notification system for terminology changes. Furthermore, SDL offers a workflow and life cycle management tool, SDL MultiTerm Workflow powered by Kaleidoscope quickTerm (SDL 2016). The product TermWeb by Interverbum Technology AB (Interverbum Technology AB 2016) also includes workflow functionalities.

4.3 Quality assurance at product level

Most QA in terminology work is concentrated on the product level. In terminology work the product is the terminological resource, in most cases the terminological database as defined in section 2. The main units of the TDB are the terminological entries, also called terminological records, containing the terminological and linguistic data related to the concepts as well as administrative information (ISO 26162:2012, ISO 22128:2008).

Since quality is not an absolute value and highly depends on the expectations and needs of all the stakeholders, the set of inherent characteristics can be mainly grouped in the following categories (cf. Chiocchetti et al. 2013a: 28ff, Lušicky and Wissik 2015: 69ff):

- linguistic criteria (e.g. linguistic correctness, appropriateness of the terms in the given context or domain, correctness of phraseology),
- content criteria (e.g. correctness of the relation between the terms in the source and target language, correctness of subject-field attribution, correctness of the definition, domain coverage),
- formal criteria (e.g. completeness of terminological entries, correctness of data field attribution, correctness of language attribution, correctness of cross-references, elementary nature of data categories).

These criteria can be checked manually, within and across entries, either by the quality evaluator or with the help of a specialised software.
Software can help identify and alert about quality issues in terminology resources (on quality issues cf. also Zorrilla-Agut 2014: 539). Quality issues can be of varying scope and of different degree of severity:

- **Scope**: Does the flaw stem from an isolated single data value within a terminological entry (e.g. misspelling in a definition)? Does it affect only some languages (e.g. English terms are of good quality, whereas other languages still require improvements)? Does it affect the terminological entry as a whole (e.g. wrong classification)? Are terminological entries questionable in their relation to other terminological entries of the resource (e.g. doublettes or gaps)?
- **Severity**: How serious is the flaw? This depends on the purpose and the target users of the resource, which determines what might be considered major or minor quality issues.

Generally, the software to improve the quality in terminology resources uses alerts. Rarely the software auto-correction itself; problems are rather spotted and then presented to the quality evaluator for correction.

### 4.3.1 Quality assurance within terminological entries

Already simple software algorithms can reduce a lot of the effort required to spot quality issues within a terminological entry (cf. Schmitz and Straub 2010: 109; Wetzel et al. 2012, 2013; Chiocchetti et al. 2013c). Some examples:

- **Spell checker**: Alert about misspellings of terms, typos, double white spaces, etc.
- **Custom filter rules**: Identify terminological entries where some desired or mandatory content is missing (such as a definition or approval status) or alert about the absence of a language in a terminological entry.
- **Custom term formation rules**: Alert, for instance, about terms that exceed a given string length.
- **Broken links**: Alert about non-resolving internal (within the resource) or external references.

These measures all help to improve consistency, editorial quality and content coverage within a terminological entry. At least some terminology management tools enforce the above while adding or editing data (cf. Steurs et al. 2015: 246). However, often low quality data is imported from external sources. In this case the software needs to inspect parts or the whole TDB (or export of the TDB) entry by entry. The identified flaws are then usually added to a to-do-list that is being worked on by data quality evaluators (e.g. Wetzel et al. 2013, Valentini 2016).
4.3.2 Quality assurance across terminological entries

While the above measures are relatively straightforward to implement, assuring quality across entries is a more demanding algorithmic challenge. Terms and concepts are evaluated against the presence of other terms and other concepts in the resource. A very typical and frequent case is the existence of concept doublettes.

(1) Canonisation: Are terms and term components that occur in several entries always formed in a consistent manner?

While in German both Curricula-Kommission and Curriculakommission may be acceptable forms, it should be guaranteed that throughout the whole resource only one variant, for example Curricula-Kommission is used consistently, and not alternatively, also in compounds such as Curriculakommissionsmitglied and Curricula-Kommissionssitzung.

(2) Doublette recognition: Are there two or more entries in the resource that denote the same concept?

Some terminology tools already alert about identical terms (“This term already exists. Do you really want to add it again?”) (cf. Schmitz and Straub 2010: 109, Steurs et al. 2015: 246). But this kind of test only checks whether a given term exists. It cannot assess whether the concept is already part of the resource. For example, the resource has an entry that describes the concept of ‘civil union’ and the German and Italian terms eheähnliche Gemeinschaft and famiglia di fatto are already stored within the database. If someone later tries to add famiglia di fatto again, an alert will pop up. However, if someone starts adding a new entry with the term convivenza more uxorio, no alert is triggered — since the strings famiglia di fatto and convivenza more uxorio are not similar at all. Therefore, simply looking for term duplicates has serious limitations. Only a systematic approach through a concept system can address this.

(3) Spotting conceptual gaps: Are there any entries missing that should be in the resource to exhaustively describe a domain?

It is important to understand that all these measures are part of a continuous process. With every addition or modification of a term or of an entry, another term or entry may be affected. Ideally, watch agents run continuously in the background and alert quality evaluators about the impact a current local change can have for other data.

4.3.3 Quality assurance throughout time

In highly regulated environments, such as finance, health care, or law, documentary evidence is a crucial process requirement. In such environments, standardised approval procedures, electronic signatures,
escalation paths, and audits may be adopted to prove quality. For instance, Title 21 of the Code of Federal Regulations of the U.S. Food and Drug Administration designates in Part 11 criteria under which the agency recognizes “electronic records, electronic signatures, and handwritten signatures executed to electronic records to be trustworthy, reliable, and generally equivalent to paper records and handwritten signatures executed on paper” (GPO 2015). This can be interpreted as: if something is not documented, it does not exist or has not happened.

For terminology resources and software this means that every single data change needs to be completely transparent and traceable. When adding a comment to an entry, approving a term, correcting the spelling of a term, etc., some aspects should always be transparent and traceable: Who made the change? When did it happen? Which changes were implemented (older value, new value)? For this, terminology software needs to be equipped with full revision tracking capabilities to document the evolution of data in form of a changelog.

4.3.4 Systematic approach through concept maps

Some of the above measures apply to rather isolated data values. Yet the semantic challenges, particularly doublette recognition or spotting conceptual gaps, can only be properly addressed through a systematic approach to maintain terminology resources, namely through concept maps. They set the individual entries into relation to each other through hierarchical broader/narrower as well as associative links. Translators using the conceptual semantic context thus better understand meanings instead of looking onto one or two isolated terminological entries. New terms, definitions or translations are drafted and approved in congruence with parent concepts, guided by related descriptions and contexts. Resource providers regain control over large collections of terms and concepts. From general to more specific – this divide-et-impera approach gets large information under control, and brings meaningful structure into the resource. Instead of scrolling through thousands of terms, users, e.g. translators, visually navigate up/down along the path of hyperonym, co-hyponyms, hyponyms and otherwise related concepts.

4.3.5 Quality assurance features in terminology software today

As stated by Steurs et al. “new tools and programs are released every day. Since the dawn of the first translation environment tools and terminology tools many things have changed” (2015: 225) and these “terminology tools have received ample attention in recent research” (2015: 227). However, little research has been done on QA in terminology tools. Steurs et al. (2015: 227ff) selected seven parameters for analysing different terminology tools. They found out that only three tools had a QA check for doublettes (SDL MultiTerm, memoQ and i-Term), but no spell check
functions, and that one tool (Wordbee) had a spell checker function, but no other QA functionalities (cf. Steurs et al. 2015: 246).

Furthermore, the EU project LISE\textsuperscript{5} — Legal Language Interoperability Services — focused, \textit{inter alia}, on quality issues in legal terminological databases and investigated the measures discussed above (e.g. Wetzel et al. 2012, 2013; Chiocchetti et al. 2013c). The project partner ESTeam, a Swedish software company, contributed linguistic tools to evaluate the above mentioned aspects within real data (cf. Zorrilla-Agut 2014). The technology is proven to be applied and the ESTeam Tools and LISE services are a sustained service (http://www.lise-termservices.eu).

4.4 Quality assurance at staff level

Terminology work requires carrying out highly specialised tasks. This often means that tasks and roles need to be allocated to a team of people with specialised competences and skills. In legal terminology work, teams may consist of linguists, translators, terminologists, translator-terminologists, revisers, legal experts, and other domain experts. Either they work together in a team on specific projects, or legal experts act as revisers and check the work done by linguists (Chiocchetti et al. 2013a). The constellation of teams in terminology work, and their skills and competences depend on the project objectives, the stakeholders and their requirements.

Terminology work in the legal domain in practice is often part of a structured translation service, and is therefore subject to requirements for translation services (e.g. ISO 17100:2015). Since we have established that processes in terminology work should also be transparent and traceable, the same principle applies related to human resources. The staff involved in terminology work have to be qualified for their respective position and role. Depending on the human resources needed and available, the staff should hold qualifications in the relevant field or adequate professional experience in the role or, in some cases, a recognized certificate of competence. Documented processes can ensure that the staff selected for certain tasks have the required competences and qualifications (cf. ISO 17100:2015). The requirements towards the individual roles, the competences needed and the responsibilities should be documented (e.g. in project requirement specifications, a manual, guidelines). Based on the documented requirements, the existing staff may be encouraged to do an honest assessment of their responsibilities and determine where they may have skill gaps and require further training. Similarly, they might assess whether their competences are being fully exploited (cf. Lušicky and Wissik 2015: 66). In addition, workflow charts can be helpful to visualise the individual roles and to enforce their ownership. As terminology is a dynamic and evolving field, continuous training for all staff involved is highly recommended and should be included in the quality management plan.
4.4.1 Terminologists

Staff with terminology-related expertise (either terminologists proper or translator-terminologists) are reported to play a central role in the planning phase, in the selection of reference material, during research and documentation of designations, and in the management and maintenance of terminological resources (cf. Chiocchetti et al. 2013a). They should be familiar with terminology theory and practical terminology work. Depending on the size and complexity of the terminological resource, language requirements, and the number of stakeholders involved, further specialisation of the role may be needed, such as senior and junior terminologists. When working in teams, terminologists usually give each other feedback and carry out internal quality checks, but QA should be a systematic process and should therefore be included in the quality management plan. Terminologists may also need to train domain experts, who act as quality evaluators, on basic terminological principles, quality issues and proper evaluation techniques. For the quality evaluators, terminologists may need to extract the relevant terminological entries, draw up guidelines and checklists that make the evaluation more straightforward. This ensures for example that domain experts concentrate on content revision instead of being distracted by correcting spelling mistakes.

4.4.2 Quality evaluators

Revisers of terminology are usually experienced terminologists or translators with a solid knowledge and skills in terminology work. In translation-oriented terminology work, revisers are often on the one hand in charge of the revision of translations, and on the other hand responsible for the revision of the terminology proper. The terminological revision involves reviewing both the form and the content of each terminological entry. The revisers may check the entry for accuracy of the equivalence, presence of a textual match in the source and target texts, accuracy of subject field attributions, sources, etc. (cf. Lušicky and Wissik 2015).

Domain experts (subject matter experts) are experts in one or more subjects that are being treated in the course of terminology work. In contrast to some other roles described in this section, the expertise of domain experts does not necessarily have a multilingual dimension. In the scope of terminology work, domain experts can act as consultants, revisers, standardisers, or — more rarely — as terminologists proper (cf. Chiocchetti et al. 2013a).

As revisers in the terminology workflow, domain experts (legal experts and other domain experts) are ideally involved in content revision. Domain experts are often not familiar with the principles of terminology work or terminology revision (Chiocchetti et al. 2013b: 14). They should be
instructed on the objective and the target users of terminology work and should ideally obtain checklists prepared by terminologists.

In the scope of legal terminology work, domain experts may also act as drafters of legislation. In this role they are obliged to be familiar with terminological principles (cf. Peruzzo 2012). For example, the Interinstitutional Agreement on common guidelines for the quality of drafting of Community legislation (1998), advises that “concepts or terminology specific to any one national legal system are to be used with care” and further stipulates “[t]he terminology used in a given act shall be consistent both internally and with acts already in force, especially in the same field.”

4.4.3 Quality managers

Quality managers play an important role in promoting and ensuring quality in legal translation and terminology services, under consideration of external factors, such as deadlines and remuneration (Prieto Ramos 2015). They may also exercise a coordination role and may be involved in activities related to coordinating translation or managing translation quality. Ideally, they are familiar with terminology work and have specific project management skills, e.g. acquisition, planning, managing processes, roles, and activities (Lušicky and Wissik 2015). It is the role of quality managers to initiate, plan, implement, monitor and control quality management procedures that are commensurate with the requirements of the relevant stakeholders of the multilingual legal TDB.

In order to monitor the use of a multilingual legal terminological database, which can be one of the indicators of product quality, the quality manager can choose between different instruments. Cerrella Bauer (2009) lists terminology management systems featuring automatic email generation in the interface of terminological entries that allows users to contact the terminology management team, statistic functions for monitoring term searches, and surveys on user satisfaction as possible feedback loops.

5 Conclusions

In order to ensure the quality of a multilingual legal TDB and in terminology work in general, it is not enough to consider only the final product – the terminological resource itself. In fact, a quality assurance framework for legal TDB should consider the QA of the following three aspects:

- persons,
- processes,
- products and services

supported by the best suited technology.
In this article we have described how such a comprehensive terminology QA framework may be set up. We have also discussed the different aspects of QA as well as currently available software features supporting terminology QA in a (semi-)automatic way. It has to be pointed out that each step “plays some role in maintaining and improving the overall quality of terminology” (de Coronado et al. 2009: 532), therefore a comprehensive approach as discussed in this paper is critical to ensuring the quality of terminological resources and services.

Bibliography


• **Chiocchetti, Elena, Wetzel, Michael and Tanja Wissik** (2013c). LISE — Investitionsschutz für gewachsene Terminologieressourcen. *eDITion* 2, 31-32.


• **de Coronado, Sherri, Wright, Lawrence W., Fragoso, Gilberto, Haber, Margaret W., Hahn-Dantona, Elizabeth A., Hartel, Francis W., Quan, Sharon L., Safran, Tracy, Thomas, Nicole and Lori Whiteman** (2009). “The NCI Thesaurus quality assurance life cycle.” *Journal of Biomedical Informatics* 42, 530-539.


detail/-/publication/53a7617f-822e-4d4b-9a0b-fe518c584f78/language-en (consulted

  http://www.interverbumtech.com/Portals/0/Admin_and_User_Manual_TWPro_V3.17-
R1A_EN.pdf (consulted 15.11.2016).


  Organization for Standardization.

  Specification of data categories and management of a Data Category Registry for


  Organization for Standardization.

• **ISO 26162:2012** (2012). *Systems to manage terminology, knowledge and content —
  Design, implementation and maintenance of terminology management systems.*

• **ISO 30042:2008** (2008). *Systems to manage terminology, knowledge and content –
  TermBase eXchange (TBX).* Geneva: International Organization for Standardization.

• **Kaleidoscope** (2015). *quickTerm - Terminology for everybody - with Workflow and
  Life-Cycle Management.* https://www.youtube.com/watch?v=UVl7GDuzlqq (consulted
  15.11.2016).


the 19th European Symposium on Languages for Special Purposes, 8-10 July 2013, University of Vienna, Vienna. Vienna: University of Vienna, 534-545.

Biographies

Elena Chiocchetti is a senior researcher at the Institute for Specialised Communication and Multilingualism of the European Academy of Bolzano (Eurac Research). She holds a degree in Translation Studies from Trieste University and obtained a PhD from Bologna University with a thesis on multilingual knowledge management in SMEs. Her research interests include comparative terminology work in the legal domain, terminology standardisation, terminology workflows, terminological databases and multilingual business communication. E-Mail: Elena.Chiocchetti@eurac.edu.

Tanja Wissik is a senior researcher at the Austrian Academy of Sciences and teaches at the University of Graz. She holds a PhD from the University of Vienna in Translation Studies. She has been working in numerous national and international research projects in the field of language resources and digital humanities. Her research interests include legal terminology, terminology management, terminology workflows, language resources, translation technology, and research infrastructures. E-mail: Tanja.Wissik@oeaw.ac.at.
Vesna Lušicky is a lecturer and research associate at the Centre for Translation Studies at the University of Vienna. She is involved in a number of European and national research projects in the field of language resources and language technology. Her research interests include translation technology, terminology management, management of language resources, and digital humanities. E-mail: vesna.lusicky@univie.ac.at.

Michael Wetzel has 20+ years experience in language technologies and associated applications such as globalisation, documentation, and content management systems as well as text mining, enterprise search, multilingual classifications and nomenclatures. He is addicted to all kinds of terminology, taxonomy, ontology, enterprise vocabulary, thesaurus and related methods to structure and leverage multilingual knowledge. Michael also works as a Senior Product Manager for ESTeam AB. E-mail: michael@coreon.com.

Notes


These are only a few examples. There are many more studies on terminology workflows. For further readings and a comprehensive overview on terminology workflows see for example Kudashev 2013.

Austrian term for collegial bodies at universities with decision-making power for enacting and changing the curricula for degree programmes and certificate university programmes for further education. (UG (Universities Act) 2002, § 25, (1) Z10, § 25, (8) Z10, official English translation, available at [https://www.ris.bka.gv.at/Dokumente/ErV/ERV_2002_1_120/ERV_2002_1_120.pdf](https://www.ris.bka.gv.at/Dokumente/ErV/ERV_2002_1_120/ERV_2002_1_120.pdf).)