

Bridging Communication Gaps between Legal Experts in Multilingual Europe: Discussion of a Tool for Exploring Terminological and Legal Knowledge Resources

Peter De Baer, Koen Kerremans, Rita Temmerman

Erasmushogeschool Brussel,
Centrum voor Vaktaal en Communicatie,
Trierstraat 84, B-1040 Brussel, Belgium

Abstract

In Europe, communication between legal experts is often hampered as a result of multilingual and cultural differences. For example, legal concepts either have different implementations in national laws or only exist in some regional legislations. Consequently, a one-to-one semantic equivalence between legal terminology in different languages is rare. To address these problems, and support and facilitate the communication between welfare professionals in a European context, we set up a small pilot project. Together with legal specialists in the area of senior care, and in line with the Termontography method, we developed a knowledge base covering trilingual terminology (English, French and Dutch), centred around the concepts 'CareHome', 'NursingHome' and 'ShortStay', in four different legal settings (UK, France, Belgium and the Netherlands). In this paper, we discuss the functionalities and the underlying model of the data-mining tool 'KBExplorer', developed for browsing through the resulting terminological and legal information resource.

1 Introduction

The impulse to start this pilot-project was given by welfare professionals discussing senior care within a European context. One of their problems concerns the fact that seniors in the European union are free to choose the country and the location where they are to receive (medical) care. Providing these seniors with accurate information about the different conditions and services in different European regions, proved to be cumbersome. One of the reasons for this is that each European country has culturally specific definitions for similar but non-equivalent phenomena and consequently one-to-one equivalence between existing terms in different languages is rare. Moreover the terminology on these aspects of welfare is embedded within a culture-specific model of laws and regulations and can only be correctly interpreted within each specific legal model. With respect to European senior care, existing multilingual glossaries and dictionaries are largely inadequate.

However, European decision makers wish to harmonise the different legislations regarding the welfare sector. Legal experts are thus in need of a tool that allows them to easily search for multilingual terminology and to compare and understand the differences between the underlying legal concepts based on relevant contextual information (linguistic context, cultural context and cognitive context) (Temmerman 2000; Temmerman 2005).

In order to better understand the requirements for such a tool, we are setting up an international pilot-project together with FOS (i.e. 'Federatie van Onafhankelijke Seniorenzorg') and ANASTE (i.e. 'Associazione Nazionale Strutture Terza Età') on two aspects of welfare: 'care for the elderly' and 'elderly care institutions'. The terminological and legal information that covers these aspects of welfare was used to construct a theoretical knowledge base model and to develop the KBExplorer tool that allows welfare professionals to explore the knowledge base in a flexible yet powerful way.

Section 1 of this paper explains the design and implementation of the experimental knowledge base model, used by the KBExplorer tool. Next, we describe a list of communication and interpretation problems that arise in multilingual and multicultural legal settings (section 2), and show how the KBExplorer tool addresses each of these problems (section 3).

2 Design of the multicultural legal knowledge base model

The essence of any knowledge base management system is its ability to store and retrieve information and to create and display the relationships between units of information. Based on this hypothesis and experience in areas such as the Semantic Web (Davis et al. 2003), document management (Pepper 2002), terminology and knowledge engineering (Mhaana 2004; Streiter et al. 2004), we selected the following three relationship types as a basis for our knowledge base model: generic-specific, whole-to-part and associative. The generic-specific and whole-to-part relations may be used to group related information together and to identify and retrieve this information by means of an abstract group name. In our model we use a concept tree to represent a hierarchy of generic-specific concepts. The whole-to-part categories are hierarchically represented as a category tree. We distinguish different types of associative relations, depending on the type of information involved. In our knowledge base model we want to be able to store and retrieve (legal) texts, text phrases or terms and (term) descriptions. We therefore implemented the following types of associative relations:

- **Concept relations:** to create a formal representation of the domain at hand by means of the domain concepts and their intercategory relationships.
- **Term relations:** which may relate terms within a concept (e.g. abbreviation-full form, preferred term etc.), to relate terms that belong to different concepts (e.g. false friends) or to relate terms that belong to different categories (e.g. different languages or regional languages).
- **Text or source relations:** hyperlinks that refer to other texts. Especially for legal texts this kind of relation is very useful since most legal texts refer to other legal texts.

Based on the advice of welfare professionals, we decided to study the following topics in a number of legal texts concerning senior care: infrastructure (more specifically: room surface), personnel (further divided into qualification of personnel and number of personnel), stay costs and care costs. For each type of legislation a category was created as part of the category tree, and used to classify the corresponding laws and legal documents.

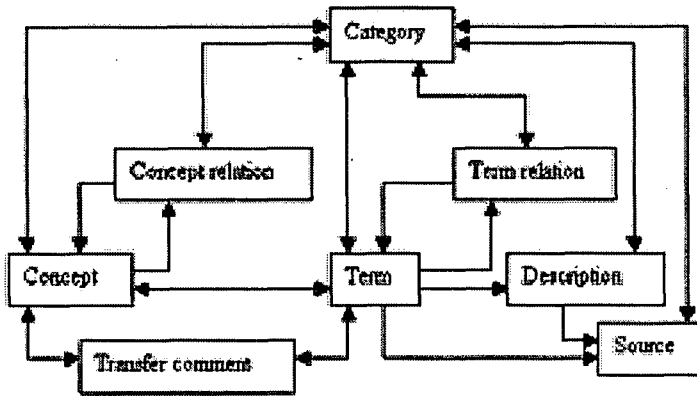


Figure 1. Knowledge base model

In Figure 1 the knowledge base model for our multilingual terminological and multicultural legal information resource is shown. An arrow indicates possible references or pointers. The flexibility of the model is obvious for most arrows are bi-directional and all information, displayed as boxes, is referenced in multiple ways. At the top of the image we see how a category may be used to classify concepts, concept relations, terms, term relations, (term) descriptions and source texts. A category in the model represents a basic property (e.g. the language of a term), which may be used to classify or label the referenced information. We noticed that, in most cases, a small number of properties is sufficient to classify a unit of information, however the potential number of properties to classify all information may be very high. Therefore we decided to use individual properties to group information, instead of the more static implementation using tables of properties in traditional database systems.

A concept represents an idea central to the domain and is used to classify terms. In our model, a concept acts as a placeholder for (near-)equivalent or (near-)synonymous terms. The terms further explain the prototypical structure of the concept. To denote the difference between the abstract concept and the term a transfer comment may be added (similar to the TermBase eXchange format (Lonsdale et al. (2002))). To allow the user to easily grasp the meaning of the concept, each concept is labelled by means of an English phrase without white spaces and with each first letter of a word capitalized (e.g. CareHome). A culture independent English description of the concept distinguishes it from the other sibling concepts. Such general concept description may use other concepts from the knowledge base. With respect to the project on senior care, welfare professionals were asked to identify the key concepts used in the domain. For this pilot-project, three concepts were selected and added to the concept tree:

1. **CareHome**: a home providing domestic care for people who are not in need of medical care except in the case of temporary illness.
2. **NursingHome**: a home providing domestic care for people who require permanent medical care.
3. **ShortStay**: domestic care for people during a short period in time.

3 Resolving legal communication and interpretation problems

When decision makers want to discuss the welfare legislation in a multicultural context, they first have to identify and be able to understand the key concepts involved in the domain. Such key concepts not necessarily apply for each region, but in order to state the differences between regions any important concept must nevertheless be defined. We could state that a shared culture-independent list of concepts is necessary as a point of reference in each discussion. KBExplorer uses a concept tree, to represent the key concepts in a hierarchical structure, which may be used to explore the domain at hand. Furthermore, the KBExplorer tool allows users to retrieve relevant information from the knowledge base by either starting from categories, concepts, sources or terms (Figure 2).

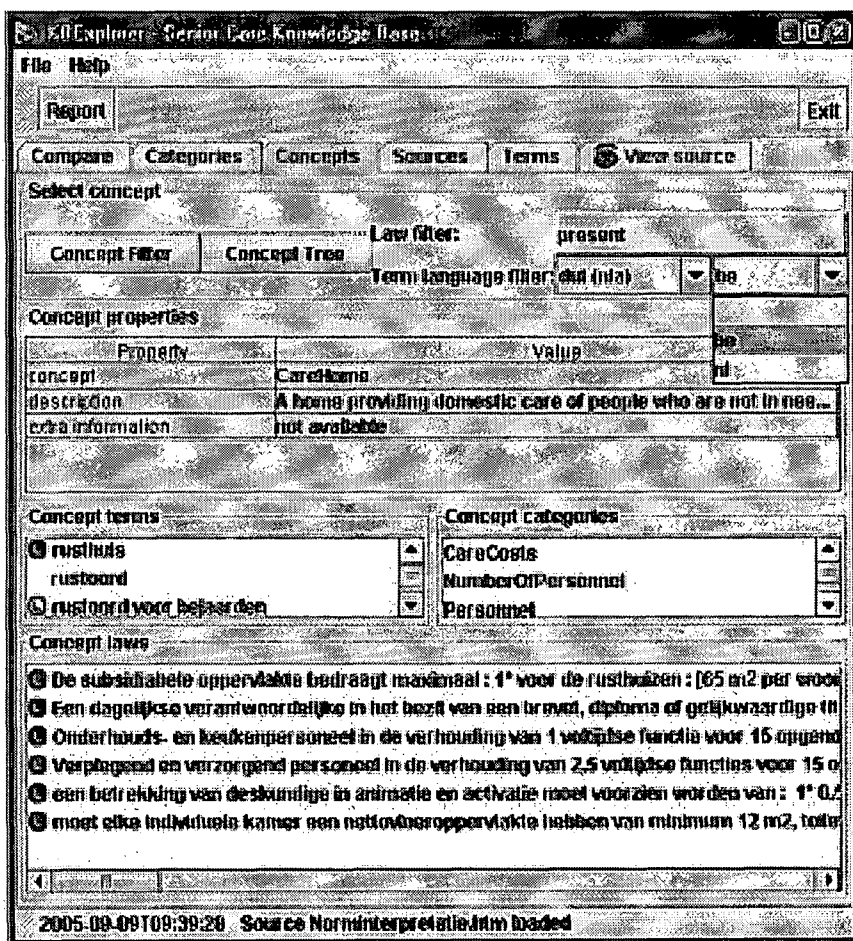


Figure 2. KBExplorer – Concepts tab

Each tab or view also offers the necessary filter criteria to further refine the information presented. For instance, the list of concept terms could be restricted to the Dutch language whereas the list of laws could be restricted to certain regions and periods of application.

Displayed information can again be used to retrieve additional information. The user could for instance double click on a law regarding the selected concept and view the legal context in the 'View source' tab.

In multilingual Europe, welfare professionals are often in need of translations of the specific terminology used in the domain at hand. KBExplorer therefore allows the knowledge base to be queried, for specific terms, using the term's language and other criteria as a filter mechanism. If for example the user selects the French term 'maison de repos', the tool will show all relevant information for this term including the corresponding concept CareHome. The data-mining tool also displays other term related information, useful for communication, like syllabification, pronunciation, etc.

To harmonise the European jurisprudence, decision makers should of course be able to compare the welfare legislation in different regions. The tool therefore provides a means of comparing law, source and term query results using different selection criteria. The query engine can handle complex category selections by using a combination of mereology and Boolean logic. A comparison between the past, the present and the future is possible based on the date enacted and date repealed of a legal source document or even an individual law.

4 Conclusion

In this paper, we presented an experimental knowledge base model to structure multilingual terminological and multicultural legal information based on three elementary relationship types, e.g. generic-specific, whole-to-part and associative. The potential usage comprises all traditional terminological products such as general and special monolingual and multilingual dictionaries, glossaries and translation memories as well as imaginable new products such as definition dictionaries or false friend lists.

We demonstrated the value of such knowledge base model by solving a list of common communication and interpretation problems in multicultural legal settings. Furthermore, we showed how the knowledge discovery tool 'KBExplorer' could be used to browse through the resulting knowledge base in a flexible and powerful way.

References

- Davis, J., Fensel, D. and Van Harmelen, F. (2003), *Towards the Semantic Web: Ontology-Driven Knowledge Management*, West Sussex, John Wiley & Sons Ltd.
- Lonsdale, D., Ding, Y., Embley, D.W. and Melby, A. (2002), 'Peppering knowledge sources with SALT: Boosting conceptual content for ontology generation', in Ide, N. and Welty, C. (eds.), *Proceedings of the AAAI Workshop: Semantic Web Meets Language Resources*, pp. 30-36.
- Mhaanna, K. (2004), 'Concept relations from OntoQuery: defining principles and application areas', *PhD Course on Language & Ontology*, Copenhagen Business School.
- Pepper, S. (2002), 'The TAO of Topic Maps. Finding the Way in the Age of Infoglut', <http://www.ontopia.net/topicmaps/materials/tao.html>
- Streiter, O., Ralli, N., Ties, I. and Voltmer, L. (2004), 'BISTRO: the online platform for terminology management', in Temmerman, R. and Knops, U. (eds.), *Linguistica Antverpiensia New Series, Spe-*

- cial Issue on The Translation of Domain Specific Languages and Multilingual Terminology Management*, Antwerp, Hogeschool Antwerpen, pp. 203-216.
- Temmerman, R. (2000), *Towards New Ways of Terminology Description. The sociocognitive approach*, Amsterdam/Philadelphia, John Benjamins.
- Temmerman, R., Kerremans, K. and Vandervoort, V. (2005), 'La Termontographie en contexte(s).' *Actes des 7ièmes Journées scientifique du Réseau de chercheurs Lexicologie, Terminologie et Traduction*.