Integrating terminological methods in a framework for collaborative development of semi-formal ontologies

Sérgio Barros[1,2], Rute Costa[1], António Lucas Soares[1,3], Manuel Silva[2,4]

[1]CLUNL - Universidade Nova de Lisboa, Avenida de Berna 26-C 1069 - 61 Lisboa;
[2]INESC Porto, Rua Dr. Roberto Frias, s/n 4200, Porto-Portugal;
[3]Department of Informatics Engineering, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, s/n 4200, Porto-Portugal;
[4]ISCAP-IPP - Rua Jaime Lopes Amorim, s/n, 4465-004 S. Mamede de Infesta

E-mail: sergio.barros@fcsh.unl.pt, rute.costa@fcsh.unl.pt, als@fe.up.pt, mdasilva@iscap.ipp.pt

Abstract (10-point Times New Roman bold, centred)

Despite the availability of tools, resources and techniques aimed at the construction of ontological artifacts, developing a shared conceptualization of a given reality still raises questions about the principles and methods that support the initial phases of conceptualization. To tackle this issue a collaborative platform was developed where terminological and knowledge representation processes support domain experts throughout a conceptualization framework.
In this article we describe the integration of a terminological method to support experts in eliciting and organizing concepts of their domain. The method is based on a linguistic analysis of textual resources with the help of a term extraction tool and by highlighting markers of relations between concepts. An application scenario is then presented to illustrate the connection between the terminological processes and the knowledge representation processes without blurring the theoretical distinction between terms and concepts.

Keywords: terminology, collaborative network, knowledge representation

1.   Introduction

An increasing number of semantic tools and resources such as concept map editors or wiki-based platforms have been built with the goal of sharing information and knowledge in collaborative networks. Despite the availability of techniques aimed at the construction of ontological artifacts, developing a shared conceptualization of a given reality still raises questions about the principles and methods that support the collaboration process. (Pereira & Soares, 2008:613) underline limitations in the development of ontologies in collaborative settings: «current knowledge about the early phases of ontology construction is insufficient to support methods and techniques for a collaborative construction of a conceptualization». Techniques may involve the (re)use of ontology design patterns (ODP), which is not without its challenges: «even users with some background on ontology modeling face difficulties when reusing ODPs for their needs» (Aguado de Cea, G. et al., 2008:45).

In the light of this issue, tasks involving conceptualization call for interplay between terminology and knowledge representation capable of rendering intuitive and operational the notions of term and concept without blurring the theoretical distinction between the different levels of analysis triggered by them. Practical work such as representing knowledge for ontology-building purposes tends to show them as alternate (sometimes opposing) sides rather than interdependent elements of a relation between objects, concepts and terms, as it is represented in the semiotic triangle in terminological science and research (e.g. Felber 1984). Considering this state of affairs, the challenge lies precisely in maintaining the premise of "terms as linguistic expressions of mental and abstract units, the concepts" throughout the conceptualization process.

In a related project – CogniNET[1] – a prototype of a collaborative tool – conceptME - is being developed to implement functionalities and models that will assist experts in the process of reaching a shared conceptualization of a given domain, in the form of semi-formal ontologies.

In this article we describe the integration of terminological methods in this tool to assist experts in the discussion and modelling of the concepts of their domain.

2.   Terminological framework

Terminology is a knowledge-related discipline whose object of study is the concept. From this perspective, since a collaborative conceptualization is developed around concepts, domain experts engaged in the collaborative process and terminologists focus on the same object. Nevertheless, while the former use terms and concepts for communicative and knowledge sharing purposes the latter study them in order to facilitate communication between experts in specialized domains or to enhance interoperability between information systems.

This twofold positioning implies that terminological methods must be accommodated to a particular communicative setting depending on an application, in this case a collaborative platform, enabling the construction of semi-formal ontologies.

To develop the work carried out in Terminology, either for

---

[1] http://cogninet.tk/

human use or machine applications, the use of texts as a resource is a common procedure. There is, nevertheless, the question of how to approach and use the text when our theoretical perspective is conceptually-based (in the line of Wüster) and the information written in the text is of linguistic origin. It is on this double dimension, linguistic and conceptual, that the method which supports the collaborative platform conceptME is based.

The platform conceptME is a technological space that allows the user to create and share conceptual systems resulting from conceptualization processes, collective or individual, which the user accepts/wants to share with a set of partners, in order to discuss and negotiate them. In these contexts, the use of natural language is unavoidable, although it carries with it, by definition, a great number of ambiguities and imprecision, characteristics that one should avoid in any negotiation process.

3. Overview of the conceptME method

The conceptualization framework in the platform is structured in four phases: concept elicitation, concept organization, concept sharing and concept discussion (Cristóvão et al., 2012). Each of these phases is supported by a set of activities related to terminology and/or knowledge representation, being that the first phase is fully supported by terminological processes, based on texts: collection, identification and classification of resources and terminological extraction. Terminological work also supports the second phase of conceptualization, when experts engage in the organization of concepts.

In terminology work, text is a relevant resource since it works as a repository that gathers linguistically structured information, from which we highlight terms and linguistic markers that play a central role in the method described in this paper. Since conceptME is aimed at domain specialists, presenting them the terms and linguistic markers that specifically occur in reference texts of their professional environment equals to offering them a key to access knowledge that, in theory, they already own.

In the following sections we describe the terminological processes that support the conceptualization phases of eliciting and organizing concepts:

i.   Analysis of textual and terminological data so as to display it in a structured way in the platform structure;

ii.   Definition of an hypothesis (an application scenario) based on structured information, that allow experts to choose the conceptualization path that better suits their needs.

4. Text: a repository of terminological information

The status and the role of specialized texts have been studied by (Costa, 2001; Costa, 2006; Costa & Silva, 2008). Specialized texts may, simultaneously, be understood as a production and a product of a restricted communication community, either professional or scientific. The text concentrates all the linguistic elements that designate and point to extra-linguistic elements that result from the interaction between language and social life, which allows one to analyze texts both as a process and as a result (Costa, 2006:80).

Terms designate concepts which in professional contexts, specific domains or for a given purpose, form conceptual systems portraying the knowledge that individuals produce and understand, in specialized texts of specific subject fields. There are, necessarily, intersections between objects, their representation and their designations. To acknowledge this triangular relation which encapsulates beliefs, scientific ideologies and a vision of the world, authors build discourses with a mono-referential value, in given contexts and for themselves. In a specialized communicative situation, authors must limit in discourse, as much as possible, the diversity of meaning constructions so as to come closer to a discourse that will ideally have one meaning, without ambiguities. Such discourses will probably never be reached and their existence is highly difficult to prove.

Given that all discursive acts (written or oral) are reflected in texts and involve complex cognitive, linguistic and social processes, a terminological and linguistic analysis of specialized texts helps to pinpoint conceptual structures behind linguist structures. As a result, when integrated in the platform, terms and markers of lexical-semantic relationships support users in their proposals of semiformal representations, thus bridging the gap between terminology and knowledge representation.

Although knowledge has an extra-linguistic nature, it is through the discourse that in most cases one is able to reach knowledge and its representations. Words are privileged means to represent knowledge. The difficulty in theorizing about it lies in the fact that those two realities – the world and its discursive representation – create a durable and reciprocal relation.

This context calls for a closer look at the description and characteristics of the specialized text as a result, i.e. a repository, as it becomes an object of observation and analysis for those who use texts to identify terms and other terminological information necessary for conceptualization. From this perspective arises the need to manage data found in texts, which in its turn, requires the management of texts as objects of knowledge, prior to analyzing their content. In view of these requirements it is necessary to create a typology of texts.

4.1 Collecting, identifying and classifying resources

When compiling a specialized corpus, one has to rigorously select a certain number of texts in the specialized domain, which will then become the objects of analysis. Such a process leads the researcher to ponder the parameters underlying the selection, organization and systematization of the texts that will constitute his/her corpus of reference.

Previous work focused on the issue of typologies (Costa, 2006), which presupposed the classification of a series of

texts organized under the same name. To that purpose texts must maintain among themselves similarity relations at the micro- and macro-structural levels through the identification of regularities which are proper to a set of texts, as opposed to regularities of another set of texts.

A typology is the result of an organization of texts based on characteristics that are common to them, which makes the classification possible. This classification allows a systematic distribution of texts in groups or types to which we attribute a label or a generic name. This grouping, which is always artificial and depends on the goals of the research and the point of view of the researcher, may take into account either linguistic or extra-linguistic factors.

A typology does not presuppose, thus, any form of hierarchy, dependency or semantic or conceptual relation between the objectives that comprise it. A typology can be built from genres or types of texts. To Maingueneu, classifying texts into types is a sociological rather than a linguistic activity, while the genre constitutes the verbal action: « Les genres de discours relèvent de divers types de discours, associés à de vastes secteurs d'activité sociale » (Maingueneau, 1998:47). For the author, constructing discourse typologies is pertinent only if you take into account the genre, founding concept of the verbal activities: « Tout texte relève d'une catégorie de discours, d'un genre de discours » (Maingueneau, 1998:45).

To talk about types of discourse means to establish parameters that are congruent with the different sectors of society, as each one of them produces discourse and texts that can be classified under a specific typology. Scientific research, for example, is a sector whose textual and discursive production constitutes a type in itself, as it constitutes the product of a specific social activity. Therefore, we think that establishing type typologies, as well as genre typologies, results from the observation of the socio-discursive conditions under which the text was produced, given the fact that it is the representative witness of a collection of texts which, in its entirety, characterizes speech.

A text corpus from a specific domain is ideally made up of texts that correspond to a typological organization with the objective of creating a certain representativeness; this representativeness is not taken in the statistical sense, but rather in the sense of texts as scientific products recognized by the members of the professional or scientific community in which and for which the text was originally written. Only with the creation of such criteria is it possible to guarantee the compliance of texts with the pre-established objectives, which are obviously the guarantee of all research work.

4.2 Towards an operable typology

Taking into account the theoretical assumptions explained above the conceptME platform integrates a typology whose goal is to allow users the organization of the texts required to extract terminological information for the purposes of a conceptualization. The typology was proposed upon the detailed analysis of texts produced in the civil construction domain, more specifically in rehabilitation.

academic text
- master dissertation
- PhD thesis
- monograph
- report

specialized publication
- journal
- dossier

legislation
- law
- decree-Law
- ordinance
- contract

technical text
- technical sheet
- technical training
- textbook
- technical report

standard
dictionary
encyclopaedia

The proposal of the categories results from the resources compiled and identified so far, that is, based on the types of documents more frequently used by the target users of cogniNET, within the rehabilitation domain. Users of the platform can increment the typology since it is an open one, in case the types already specified don't suit the users' needs. In addition, users can select a more generic type in case the more specific one is not suitable to their needs. For example, a user may not know which category suits a given text but still knows that it belongs to the 'Legislation' category. Additionally, the possibility of conceptualizing via reference linguistic resources was also considered, namely dictionaries and encyclopedias.

This typology conforms to a repository where users can organize texts of their choosing into categories, thus allowing the compilation of a customized reference corpus. Such a corpus will be dynamic and up to date at all times.

5. Terminology extraction: a different goal

The semi-automatic treatment of corpus regards the process of terminological extraction as an initial step towards the elicitation of concepts. During this phase of conceptualization domain experts can use a terminological extraction functionality which allows them to obtain from a text or group of texts a list of linguistic units that potentially designate concepts. This functionality allows the selection of one or several of these suggested candidate terms with which concepts can be organized in the following phase of conceptualization. In the beginning of the 90s, following the rapid development of computational linguistics and the widespread availability of corpora, terminology extraction became an important research interest as a

means to reduce time and effort in different tasks related to different goals. (Cabré et al. 2001:53) identify several of the goals behind terminology extraction: «building of glossaries, vocabularies and terminological dictionaries; text indexing; automatic translation; building of knowledge databases; construction of hypertext systems; construction of expert systems and corpus analysis». The task of reaching a shared conceptualization in a collaborative framework can also benefit from the potentialities of these tools. When considered in terms of such a goal, it matters to reflect on the implications that a term extraction output has for a conceptualization phase that will be carried out by individuals who have a high level of knowledge in specific domain areas, thus, capable of identifying terms and concepts without necessarily making a difference between the linguistic and the conceptual level. The challenge behind the terminological extraction is to provide to experts terminological information which serve as a starting point for their conceptualization.

5.1 Criteria for selecting a tool

After reviewing a set of existing term extraction systems, three of them were selected for an evaluation: multiwords, TermoStat and GaleXtract. The first makes use of statistical methods and the other two use a hybrid method with the incorporation of a tagger with rules of the Portuguese language.

The terminology extraction methods are usually defined by linguistic and/or statistical criteria, which accounts for the linguistic dimension of terms. The possibility of extracting a set of linguistic units based on their frequency in connection with the recognition of language patterns typical of specific languages conforms to the main goal behind the evaluation of the extractors. Moreover, it bears also a connection with the requirements of an initial conceptualization activity: to obtain a list of acceptable linguistic units. In the light of these criteria hybrid methods of terminology extraction seem the most adequate for the platform: "Statistical approaches, like the linguistic ones, used alone only seldom reach truly satisfying results" (Pazienza et al. 2005:259). Furthermore, an extractor capable of accounting for several languages is preferable to a language-independent tool.

5.2 GaleXtract: description

Based on the evaluation criteria described above GaleXtract[2] was selected as the term extraction tool since it is based on a hybrid method, and it handles several languages: Galician, Spanish, English, French, Portuguese. Its extraction allows the use of either Freeling or Treetagger for the tagging phase. Furthermore, five statistical measures can be employed although only one is available in the collaborative platform[3].

[2] Developed under the Gari-Coter project: http://gramatica.usc.es/proxectos/Gari-Coter/?lang=gl.
[3] Although the measures of coocurrences, loglike,

6. Integration of GaleXtract in the platform

The terminological extraction process consists in automatically extracting term candidates from a single text or group of texts and then select one or several to work with.
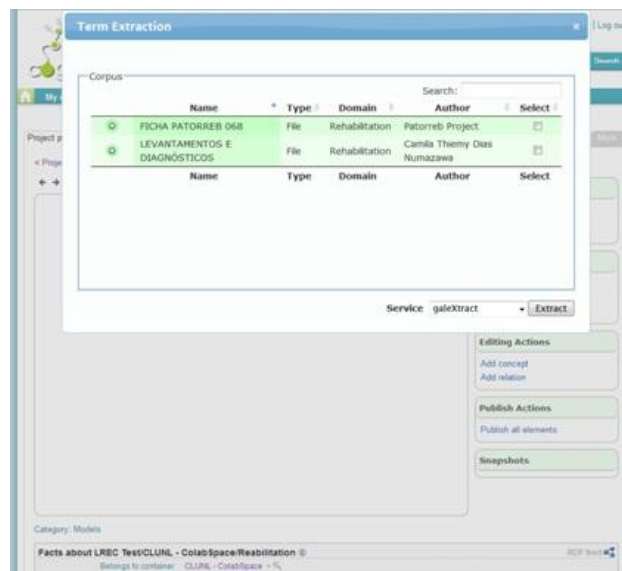

Figure 1: Term extraction from resource(s)

The output list that is presented to the expert can be sorted alphabetically or by ordering the results from the highest to the lowest statistically measured linguistic unit.
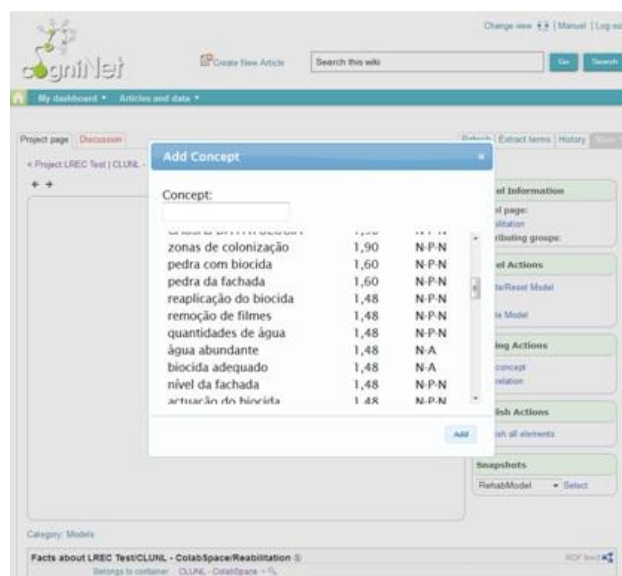

Figure 2: Term extraction result

chi-square, mutual information and scp generated similar results chi-square was the measure chosen.

Nevertheless, tools are only a means to save time and effort in terminology work: «Terminology extraction implies, almost invariably, that whatever is initially collated is a collation of candidate terms» (Ahmad 1998:141).

From a terminologist's stance, the initial selection of acceptable term candidates follows linguistic criteria, by selecting combinations of words that match patterns that are typical of the Portuguese language: noun, noun + adjective, noun + preposition + noun, i.e. *colonização biológica* (*biological colonization*). Or by selecting combinations of words whose meaning is not the result of the sum of its parts, i.e. *filmes negros* (*dark spots*).

intensidade nas zonas
filmes negros
colonização biológica
PALAVRAS-CHAVE Parede&Exterior
DESCRIÇÃO DA PATOLOGIA
CAUSAS DA PATOLOGIA
rocha total
microscopia electrónica
certa regularidade
sua aderência
raios X
Barroso de Aguiar

Table 1: Sample of extraction.

One obstacle resulting from the semi-automatic method is that the output may not conform to the needs of experts, such as combinations of linguistic units that don't designate a concept or a conceptual unit, i.e. *certain regularity* (*certa regularidade*).

Although term extraction systems are useful to obtain lists of terms, a crucial methodological step consists in resorting back to their natural habitat, texts. Furthermore, since the terminological extraction process establishes a connection with the conceptualization phase where experts organize concepts something else is required in order to guide them towards relations between concepts.

6.1 Term candidates and lexical-semantic relations

Specialized texts are undoubtedly a vehicle of knowledge. In terminology, terms play a fundamental role as nuclear elements of lexical and semantic relations that language professionals or experts are able to recognize in texts. Such relations, held between the meanings of words, form the basis for the construction of semantic networks and allow the representation of the knowledge available in a text or set of texts.

Within the conceptualization framework designed for the conceptME platform the notion of knowledge representation covers several activities, namely the identification and selection of relations, the identification and selection of terms, the representation and consistency check of conceptual structures.

Recalling the motivations behind this research, the focus of integrating terminological methods in the platform is to establish a suitable and operable connection between the terminological processes and the knowledge representation processes as a means to support experts in the organization, sharing and discussion of concepts.

Considering the principles behind the terminological processes and how such principles relate to the knowledge representation processes involved in the platform, identifying potential terms during the concept elicitation phase must be complemented with a technique/method that allows one to understand not only how a given term is used but also the relation that it can have with other terms/concepts.

Since concepts can be expressed through linguistic forms, specialized texts are valuable sources of information for terminologists carrying out tasks related to concept analysis, like semi-automatic extraction of terminology or of relations between concepts.

Domain experts will also use specialized texts – previously selected by terminologists or by themselves – as a source of knowledge for their conceptualization tasks. Therefore, a natural step in our approach is to consider contexts as a source of information about concepts and about relations between concepts.

Following the work of (Hearst, 1992) several researchers developed the idea of extracting from texts linguistic patterns that express information about concepts, as contexts from a corpus of urban rehabilitation exemplify. For example, the structure is a typically expresses a relation between a subordinate concept and a superordinate concept:

> «A pre-dosed industrial mortar **is a** mortar whose components are dosed in the factory and supplied to the construction site, where they will be mixed according to instructions and conditions of the manufacturer»

The structure *is composed of* points to a partitive relation:

> «The floating floor **is composed of** laminated wooden boards arranged in opposite layers, so as to reduce the movement of the timber.»

The structure *X is caused by Y* expresses a relation between an effect and a cause:

> «The moisture is usually **caused by** the inadequate protection of the outer wall with respect to the atmospheric conditions to which it is subjected.»

Applications of this type include the writing of definitions (Pearson, 1998), concept analysis (Meyer, 2001), semi-automatic ontology building (Gillam, Tariq, & Ahmad, 2005) or the reuse of ontology design patterns (Aguado de Cea, Gómez-Pérez, Montiel-Ponsoda, & Suárez-Figueroa, 2008).

Based on the hypothesis that contexts such as these provide useful input to those who engage in a conceptualization process an application scenario related to the domain of civil construction, specifically rehabilitation, exemplifies how this terminological data can be applied.

7. Scenarios: an application

Considering the theoretical principles described above plus the criteria behind the terminological approach to the elicitation of concepts and the support to the concept organization phase, the integration of a terminological method in the platform is illustrated below.

A scenario implies starting a conceptualization with input, which consists of term candidates manually selected from the term extraction process, complemented with contexts with information about concepts, evidenced by the presence of linguistic markers.

The first part of the application scenario draws on the first phase of the conceptualization framework, whose goal is to elicit concepts from textual resources.

patologia
colonização biológica
liquens
musgos
filmes negros
humidade
microscopia electrónica de varrimento
MEV
difracção de raios X
DRX
rocha total
biocida
nebulização

Table 2: Terminological input.

When compared with the raw output of the extraction tool these linguistic units illustrate the advantage of having a list of candidate terms that will serve to select designations of concepts to be organized in the following conceptualization phase.

The table above presents candidate terms which can trigger a conceptualization, i.e. a generic relation between the concept of biological colonization (*colonização biológica*) and that of pathology (*patologia*), a causal relation between the concept of humidity (*humidade*) and that of biological colonization. Relations such as these can be established by domain experts using a catalogue of concept relations that is available in the platform as a resource.

Figure 3: Possible conceptual relation(s)
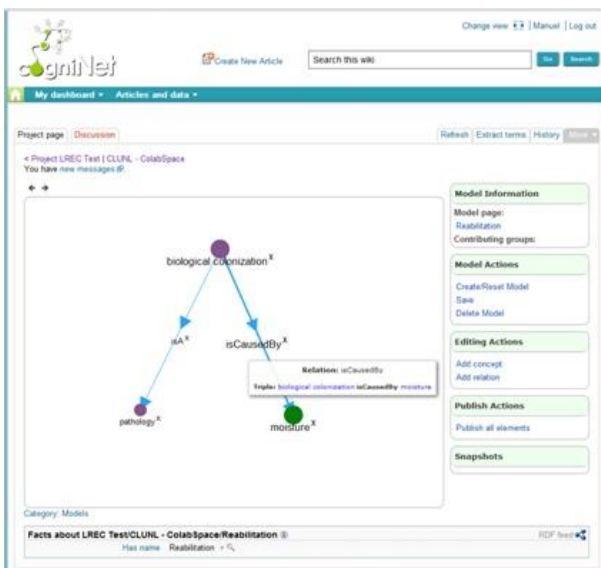
As a means to support the concept organization phase, more specifically the use of the catalogue of concept relations, experts can consult contexts where those terms occur, thus obtaining further information about the respective concepts.

The objective behind the contexts, previously selected and filtered by a terminologist, is to call the attention of experts to the presence of markers of concept relations within contexts, which helps them to decide which type of conceptual relations exist between certain concepts.

Below we present a context with a linguistic marker of cause-effect relation between the concepts of biological colonization and moisture:

> «The biological colonization of the surface of the stone facade was mainly **due to** the presence of moisture, and there has been a greater intensity in areas where run-off are larger and darker on the front (north).»

Despite the potentialities of linguistic markers, research in the field of terminology has shown that their reliability is limited by factors such as their degree of dependency to the corpus (Meyer 2001, Condamines, 2002), their portability across different domains (Marshman, L'Homme, & Surtees, 2008) or the presence of uncertainty markers (Marshman, 2008).

An interesting example of such limitations is provided by the following context:

> «Darkening and wood stains **caused by** the presence of moisture and staining fungi most often located at the bottom of the door, **due to** lack of inclination of the sill with the accumulation of a water layer which penetrates inside the wood.»

The context above should give rise to causal relations such as the one between the concepts of moisture and that of wood stains or between the concepts of staining fungi and that of wood stains.

Figure 4: Possible conceptual relations

In addition to the relations modeled above this context is particularly interesting for a distinction between the markers *caused by* and *due to*, both causal but in principle expressing different types of causality that only experts can recognize. The marker *caused by* refers in principle to a causal agent of darkening and wood stains (*moisture, staining fungi*) and the marker *due to* possibly refers to its explanatory cause (*lack of inclination of the sill*).

Several authors studied the nature and number of concept relations (Feliu, 2004; Nuopponen, 2005, 2011; Sager, 1990). For example, (Nuopponen, 2011) has devised a model for cause-effect relations where she distinguishes various types of causes and of effects. Around the core concept of effect the author underlines different relations, i.e. a patient relation, a symptom relation, a consequence relation, a counteraction relation and a cause-effect relation. She also sees three types of effects (resulting product, resulting state, resulting event) and different possible causes (causal agent, producing cause, explanatory cause) (cf. Nuopponen 2011:12).

The author's perspective is: «Causal relation is often seen as a relation between the concepts of cause and effect (causal sequence), but this is only the basis for a complex concept system that is often involved» (Nuopponen, 2011:12). Some authors suggest that it is not very practical to have a very detailed account of concept relations (Madsen, Pedersen, & Thomsen, 2001:7). However, if the purpose is to negotiate meaning and clarify concepts then it may be a good idea to have a breakdown of the most general conceptual relations into more detailed ones such as the ones that (Nuopponen, 2011) proposes in her causality model.

From a terminologist's perspective it would be interesting to see whether users recognize and discuss the meaning of different markers of causal relations such as the ones that occur in the context above.

s of causal relations such as the ones that occur in the context above.

8. Concluding remarks

This article described the integration of a terminological method in a collaborative framework to assist the domain expert throughout the initial phases of a conceptualization process. More specifically, we focused on the integration of a tool in the platform to extract term candidates and on supporting the use of a catalogue of conceptual relations that will be available in the platform. The organization, share and discussion of concepts is supported by natural language, more specifically texts that provide the terms to designate concepts or the linguistic mechanisms to establish relations between concepts. Nevertheless, those texts contain ambiguities and uncertainties that experts may not recognize.

The hypothesis behind this method is that eliciting concepts from textual resources and identifying concept relations for conceptualization purposes can benefit from an approach that maintains a distinction between terms
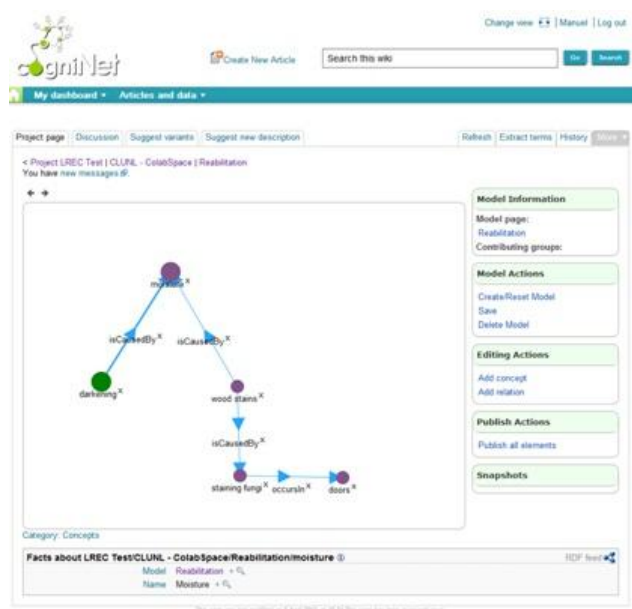
and concepts throughout the conceptualization process.

To obtain insights on the usability of the terminological method in the platform several scenarios with terminological data were prepared as application examples. Scenarios such as these are important not only to obtain an empirical insight on the connection between terminology and knowledge representation but also on the usefulness of contexts with markers of conceptual relations as a means to complement the use of the catalogue.

9. Acknowledgements

10. References

Aguado de Cea, G., Gómez-Pérez, A., Montiel-Ponsoda, E., & Suárez-Figueroa, M. C. (2008). Natural Language-Based Approach for Helping in the Reuse of Ontology Design Patterns. In A. Gangemi & J. Euzenat (Eds.), *Knowledge Engineering: Practice and Patterns* (Vol. 5268). Springer Berlin / Heidelberg, pp. 32-47.

Ahmad, K. (1998). Specialist Texts and their Quirks. In *TAMA '98: Proceedings of the 4th TermNet Symposium*. Vienna, Austria: TermNet, pp. 141-157.

Cabré, M. T., Estopà, R., & Vivaldi, J. (2001). Automatic term detection. A review of current systems. In D. Bourigault, C. Jacquemin, & M.-C. L'Homme (Eds.), *Recent Advances in Computational Terminology*. Amsterdam; Philadelphia: John Benjamins Publishing Company, pp. 53-87.

Condamines, A. (2002). Corpus Analysis and Conceptual Relation Patterns. *Terminology*, 8(1), pp. 141-162.

Costa, R. (2001). Pressupostos teóricos e metodológicos para a extracção automática de unidades terminológicas multilexémicas. PhD Thesis. Lisboa: Universidade de Lisboa.

Costa, R. (2006). Corpus de spécialité : une question de types ou de genres. In H. Béjoint & F. Maniez (Eds.), *De la mesure dans les terme. Hommage à Philippe Thoiron*. Lyon : PUL, pp. 313-324.

Costa, R. & Silva, R. (2008). De la typologie à l'ontologie de texte. In *Terminologie & Ontologies : Théories et Applications. Actes de la deuxième conférence - Toth Annecy – 2008*. Annecy : Institut Porphyre. Savoie et Connaissance, pp. 3- 16.

Felber, H. (1984). M*anuel de terminologie*. Paris: United Nations Educational, Scientific and Cultural Organization - International Information Centre for Terminology.

Feliu, J. (2004). Relacions conceptuals i terminologia: anàlisi i proposta de detecció semiautomática. PhD Thesis.

Gillam, L., Tariq, M., & Ahmad, K. (2005). Terminology and the construction of ontology. *Terminology*, 11(1), pp. 55-81.

Hearst, M. A. (1992). Automatic Acquisition of Hyponyms from Large Text Corpora. In *Proceedings of the Fourteenth International Conference on Computational Linguistics*. Nantes : France, pp. 539-545.

ISO 704:2000 - Terminology work - Principles and methods.

Madsen, B. N., Pedersen, B. S., & Thomsen, H. E. (2001). Defining Semantic Relations for OntoQuery. In P. A. Jensen & P. Skadhauge (Eds.), *Proceedings of the First International OntoQuery Workshop, Ontology-based interpretation of NP's*. Kolding: Department of Business Communication and Information Science, pp. 57-88).

Maingueneau, D. (1998). *Analyser les textes de communication*. Paris: Dunod.

Marshman, E. (2008). Expressions of uncertainty in candidate knowledge-rich contexts: A comparison in English and French specialized texts. *Terminology*, 14(1), pp. 124-151.

Marshman, E., L'Homme, M.-C., & Surtees, V. (2008). Portability of cause–effect relation markers across specialised domains and text genres: a comparative evaluation. *Corpora*, 3(2), pp. 141-172.

Meyer, I. (2001). Extracting knowledge-rich contexts for terminography - A conceptual and methodological framework. In D. Bourigault, C. Jacquemin, & M.-C. L'Homme (Eds.), *Recent Advances in Computational Terminology*. Amsterdam; Philadelphia : John Benjamins Publishing Company, pp. 279-302.

Nuopponen, A. (2005). Concept relations – An update of a concept relation classification. In B. N. Madsen & H. E. Thomsen (Eds.), *Proceedings of TKE 2005 – 7th International Conference on Terminology and Knowledge Engineering*. Copenhagen, pp. 127-138.

Nuopponen, A. (2011). Methods of concept analysis – tools for systematic concept analysis. Part 3 of 3. *LSP Journal*, 2(1), pp. 4-15.

Pazienza, M. T., Pennacchiotti, M., & Zanzotto, F. M. (2005). Terminology Extraction: an analysis of linguistic and statistical approaches. In S. Sirmakessis (Ed.), *Knowledge Mining*, Vol. 185, pp. 255-279.

Pearson, J. (1998). *Terms in Context*. John Benjamins Publishing Company.

Pereira, C., & Soares, A. L. (2008). Ontology Development in Collaborative Networks as a Process of Social Construction of Meaning. In R. Meersman, Z. Tari, & P. Herrero (Eds.), *On the Move to Meaningful Internet Systems: OTM 2008 Workshops*, Vol. 5333. Springer Berlin / Heidelberg, pp. 605-614.

Sager, J. C. (1990). *A practical course in terminology processing*. Amsterdam/Philadelphia: John Benjamins Publishing Company.