

# INTEGRATING SEMANTIC RESOURCES TO SUPPORT SME KNOWLEDGE COMMUNITIES

Manuel Moreira Silva<sup>13</sup>, António Lucas Soares<sup>12</sup>, Dora Simões<sup>14</sup>

{mdasilva@iscap.ipp.pt, als@fe.up.pt, dsp@isca.ua.pt}

<sup>1</sup>INESC Porto, Rua Roberto Frias, Campus da FEUP, 4200 Porto Portugal

<sup>2</sup>University of Porto, Faculty of Engineering, Rua Roberto Frias s/n, 4200 Porto Portugal

<sup>3</sup>ISCAP, IPP, Rua Jaime Lopes Amorim, S. Mamede Infesta Portugal

<sup>4</sup>University of Aveiro, ISCAA, Aveiro Portugal

The development of ontologies to unify and to put into context the different concepts and terms of the sometimes rather traditional and locally coloured construction industry domains is a necessary step to avoid misinterpretations and inefficient communication. The KNOW-CONSTRUCT project, as an approach to this task, decided to re-use, as far as possible, existing ontologies, classification systems and other semantic resources in order to develop a system for the integration, management and reuse of the area specific knowledge via a common knowledge base in order to consolidate and provide access to integrated knowledge, making community emergent knowledge a significant added value.  
*Copyright © 2006 IFAC*

Keywords: ontologies integration, knowledge representation, knowledge communities.

## 1. INTRODUCTION

It is well known that organizations and the environment in which they operate have considerably changed in the last few decades. The particular case of organizations in the construction sector is no exception to the rule, as they have to renew themselves more rapidly to adapt to a more competitive and changing environment, be more flexible than in the past and manage their knowledge assets in more sophisticated ways. Most knowledge management systems have emerged from document-centric approaches and are able to support, even if in a very efficient manner, only a fraction of the whole knowledge cycle (classifying, storing, and retrieving knowledge).

The Know-Construct project<sup>1</sup> intends to improve the effectiveness of the Construction Industry (CI) SME's by improving and extending the relationship with their customers through an innovative support regarding information and knowledge about

products, processes and associated issues. This is achieved through specifically developed tools, supporting in particular the formation and operation of SMEs knowledge communities in the context of Industry Association Groups (IAG). Its most specific objectives are:

- 1 To develop a platform to support the formation and management of a community of CI SME's, coordinated by an association, fostering collaboration and knowledge sharing among the members. Knowledge to be shared includes, besides product and services information, companies' experience (e.g. best practices). This will lead to a wider and deeper technical and professional competence shared by the SME's community, fundamental to the satisfaction of customer needs, obtained through closer co-operation and knowledge exchange.
- 2 To provide problem-solving support to the customers of individual IAG members regarding the selection of products, their applications and processes, as well as addressing other related problems such as legislative issues, safety issues etc. This will be materialized in an internet-based platform that will offer the possibility to establish a "one to one" system, manufacturer (wholesaler) SME to customer, advising on specific topics, although relying on knowledge created and maintained by a community of SME's mentioned in the previous point.

---

<sup>1</sup>COLL-CT-2004-500276 KNOW-CONSTRUCT Internet Platform for Knowledge-based Customer Needs Management and Collaboration among SMEs in Construction Industry (2005-2007). Project co-funded by the European Community under the "Horizontal Research Activities Involving SMEs - Collective Research" Programme.

KC system was designed to provide comprehensive services to its users regarding a large scope of construction issues, although centred in the SME community knowledge. As such, whenever a consultancy request is not completely answered by the information existing in the community, KC system will search, collect and integrate information from external sources to complement local knowledge.

## 2. SME KNOWLEDGE COMMUNITY SUPPORT SYSTEM

The relationship between organizations and communities, including their implication in the organization structure, IT systems and business model, depends on the nature of the value derived from them. This area has not been sufficiently mapped before, but (Cornejo 2003) explores it in depth starting with the nature of knowledge, the translation into value for individuals and organizations, the definition of community taxonomies and its influence on organizational drivers. KC addresses this specific point: how to obtain the maximum value to individual SMEs from a knowledge community in which those companies participate.

### 2.1 Models for knowledge communities

KC's general objective is to improve the relationship of CI SME's with their customers by offering the later an innovative support regarding information and knowledge about products, processes and associated issues. This is achieved through specifically developed ICT tools, in particular tools that support the formation and operation of SME's knowledge communities, fostering an improved collaboration aimed at generating broader and more accurate knowledge to be used in satisfying customers.

The broader context of this development are the sector's associations that provide, besides others, some sense of grouping to the participant SMEs. This knowledge community (Construction Industry Knowledge (CIK) community) can be considered as a hybrid community of practice and interest. Members - company employees - as individuals should see a direct utility to their particular jobs when participating in the CIK community. This utility comes into light when an employee (and consequently the company) realizes that, when solving a problem for an important customer (using the CNM system), the information/knowledge used to reach the solution comes also from the contribution of the other community members.

Nevertheless, not all the activities can be tracked as a causal benefit to the SME. For example, some chatting between two employees exchanging specific professional experiences, a report in a news, or a blog item by another employee of a concern

regarding the performance of a construction material, are activities that make sense in a community but cannot be assigned to a concrete immediate value for the organization.

Based on work of de Vries *et al.* 2004, the CIK Community can be characterized as follows:

- The goal is to develop and explore knowledge on construction area.
- There is continuous interaction between participants in order to meet these goals.
- Information and communication processes are continuously made explicit.
- It adds value to the participants (professionals and customers).
- The online meeting place is usable.
- The culture focuses on the needs of the participants as the route to high performance; involvement and participation create a sense of responsibility and ownership and, hence, greater commitment to the community.
- The context is highly complex and very unstable, and CIK Community will have to continuously comply with the expectations of its participants and their context of use.

The main question here is how to implement and make use of online knowledge communities in order to meet these goals and expectations. As an answer to this problem, we propose to adopt and adapt the reference model developed by de Vries *et al.* (2004) that describes factors that affect the implementation and use of an online knowledge community as a type of networked organizational communication.

Based on this explanation, CIK Community consists of three elements: professionals, on-line professional meeting place, and organizational context. We see the implementation of CIK Community as a continuing communication process, a constant search for a fit between these three elements:

- **Professionals** - all the professionals of the construction area and belonging to the associations of the project partnership.
- **On-line Professional Meeting Place** - the online space where the members of the community will take their discussions, debates, conversations.
- **Organizational context** - related to the organizational goals, culture, technology, etc.

### 2.2 KCS system's general architecture

As mentioned before, the KC project has a very specific goal: to enable individual SMEs to best solve the problems of their customers. Therefore, KCS is focused on pursuing this goal in the first place. Although a knowledge community encompasses, as stated before, mechanisms that surpass this simple instrumental goal, the initial vision of KCS is developed from this instrumentality. This means that KCS supports CIK community building in a broad sense, though focused fundamentally in generating broader and deeper knowledge to be used in managing the SMEs customers' relationship, particularly in problem solving.

The system (KCS) provides the following three general functions:

1. **Community building tools:** this part of KCS supports the processes of community building by providing the instruments to foster professional interaction and socialization. Forums and weblogs are two of such instruments and are tailored in KCS to be strongly integrated with the semantic structure supporting knowledge management in KC.
2. **Semantic resources management:** this is a set of infrastructural functionalities that support information and knowledge acquisition, organization and storage in KCS (and also CNM). More specifically, they enable the (i) management of classifications, thesauri and vocabulary, (ii) the acquisition of knowledge from digital content (including forums and weblogs entries, web pages, etc.) both internal to the CIK and from external sources, (iii) the maintenance of an ontology which is the base of knowledge representation, access and storage.
3. **Knowledge resources access:** creating, searching and updating knowledge resources will be a fundamental set of functionalities in KCS. Although much of the community's information/knowledge will be created in communication/interaction processes (forums, weblogs), there will also be the need to create/access knowledge in a more structured way. Digital content management and document management are the natural approaches regarding this issue.

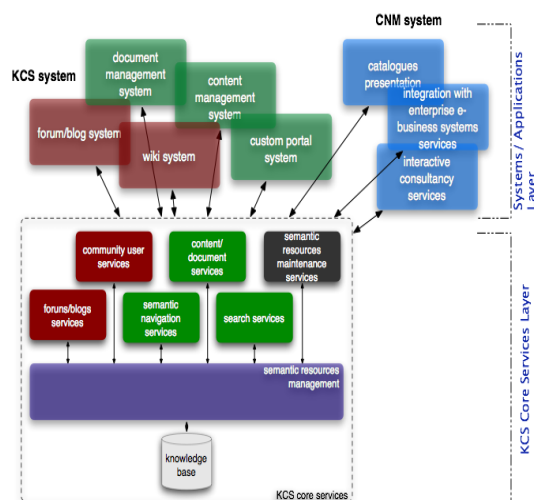


Fig. 1. Know-construct functional architecture

### 3. STUDY AND ANALYSIS OF KNOWLEDGE SOURCES FOR KCS IN THE CI SECTOR

The identification and selection of existing knowledge sources was the first step to building the semantic resources structure of KCS system. Multiple different sources like terminologies, ontologies, international classifications, standards, norms and regulations, national classifications were analysed. The difficulty of this task is well known,

since the different sources are usually designed using different theoretical grounds, design principles and serve different purposes.

#### 3.1 Types and characteristics of semantic resources

According to ISO 12006-2 (ISO, 2001) the most widely used classifications are work sections (mainly for specifications) and elements (mainly for cost analysis). They are also the most varied, not only in their itemization and structure but also in the range of other purposes to which they are put. As a result of the research, several other classifications were identified, potentially just as important, which have not yet been used to the same degree, e.g. construction products and properties/characteristics.

KC project is, along with the necessary development of classifications and taxonomies that answer the project needs, re-using/integrating as far as possible existing ontologies, classification systems and terminologies in order to develop a system that may, in the future, contribute to standards. The initial interaction of the KC project with standards issues will be to assure full compliance of the developed solution components with the current legal and de-facto standards in the targeted building sector and in relevant ICT domains.

The ontologies were developed in the areas of product characteristics, product applications and related consultancy services. These ontologies are crucial not only for the decision making support system, but also to create uniform models for customer's access. Standard or integrating ontologies do not exist in these areas. Furthermore, another essential innovation, potentially contributing to Construction Industry standards, is the development of integrating ontologies both in the areas referred above and in inter-enterprise interoperability. It is exploiting proposals used or in use in other European and international projects, in an attempt of harmonization with the current well established standards, but also as a way to oppose the unpredictable perennity of the stored data.

Besides CI online sites and other sources, some of the most relevant ontologies and classifications identified so far are: *e-Cognos ontology*; *e-Construct ontology*; *EPIC* – European Product Information Co-Operation; *UNICLASS* – Unified Classification for the Construction Industry; *IFC Model* – Industry Foundation Classes; *ICIS LexiCon*.

The identification of these sources leads to the conclusion that part of the existing information has some common principles and structures, mostly because they result from European or governmental projects which also aim to contribute to harmonization and standardization. But, its diversity, nevertheless, puts us before the problem of how to test and adapt the selected resources according to the KC project purposes and scope and the industry consortium predefined needs, taking also into consideration the specific cultural and professional

context of the ontology's development and use, the target-audience(s) and the previously defined scenarios.

### 3.2 Evaluation and selection of candidate resources

In order to choose the knowledge sources, a complex set of multi-criteria referring to different aspects were established. An analysis framework was elaborated in order to evaluate the candidate knowledge sources in what concerns their suitability to the KC system and methodology. The following table, inspired in the works of Pinto and Martins (2001) and Lelkin (2004), describes the established criteria and their scope.

Table 1. Knowledge sources evaluation criteria

<i>Know-Construct</i>		
<i>S</i>	<i>Common Criteria</i>	<i>Description</i>
All	Origin	developer(s) type of entity (CEN, ISO, DIN)
	Relevance	for the pre-defined areas of analysis
	Adequacy	for specific cultural and from domain expert point of view from ontologist point of view
	Completeness	explicit in-depth coverage
	Comprehensiveness	domains addressed in the area
	Ease of data acquisition	possibility of access and reuse (merge/integrate) language(s) availability
	Language	multilingual features
	Current status	language independence finished, work in progress, in revision
<i>Specific Criteria</i>		
Ontologies	Conceptual framework/model	ontology assumptions ontological commitment
	Type of concept	identification of generic/domain concepts and relationships
	Design principles	internal structure
	Knowledge acquisition	quality of knowledge sources adequacy of knowledge acquisition practices
	Supported applications	applications supporting the ontology codification language
	Documentation available	type of documentation available and accessibility
	Consistency	consistency of the application of the relations
Modularity	which concepts are represented in which modules	
Terminologies	Terminology purpose and scope	operational terms – functions the terminology is intended to serve
	Standardized/non-standardized	implemented as standard, other type
	Granularity	level of complexity of the available data
	Quality of the definitions	do they follow unified patterns, are simple, clear, concise, etc.
	Interconnectivity	to what extent is the terminology mappable to coding systems
	Precision and recall	retrieval effectiveness
	Normalization	of content and semantics
Responsiveness	frequency of update	
Classifications	Classification purposes	classification purposes and their relation to KC objectives
	Conceptual framework	classification assumptions and their relation to KC objectives
	Classificat. scope	domain(s)
	Type of concepts	degree of abstraction/specificity
	Previous use	use in ontology projects and outcome analysis

The choice of the evaluation methods to use depends on different factors such as:

- previously identified elements and indicators,
- knowledge to be represented,
- stage of system development,
- time and resources available,
- type of output required,
- precision / reliability desired.

Along with these criteria, subject field specialists were consulted to analyse both the methodologies and the resulting ontology. After this identification and evaluation process, and taking in consideration the general view expressed in the CWA 15142 – European eConstruction Ontology (EeO), from the candidate sources, those which, at the moment, present the best solutions for the purpose of reuse and integration of information and for the development of Know-Construct high-level ontology are: LexiCon, bcBuildingDefinitions, eCognos, IFC model and ISO 12006 – 2.

## 4. ONTOLOGIES STRUCTURE AND CONTENT

After the activities described above, the adequate strategy was delineated in order to support the structuring, maintenance and evolution of CIK ontology and local ontologies. Thus, the methodology to be followed aims mainly at identifying, evaluating and reusing existing semantic resources, like the ontologies in the IC area, due to the advantages their reuse offers: simpler design, more reliable knowledge sharing and clearer semantic representations. During the analysis described above, a perception gained strength - to deal with the very concrete reality of the SMEs of each country and with each Construction Industry Knowledge (CIK) Community of KCS, the KC system would need to have a local ontology that would answer the KC partners' particular professional and cultural needs and attend to its social context of use. This meant the need to construct not one, but more than one ontology.

KC has decided to take this fact into account and look at this possibility as part of the standardised way to develop ontologies in the sector, but in such a way as to keep a common central ontological content (structure, attributes, relations, etc.) from where to derive the more specific ontologies. The development of this more specific ontology will, thus, be based on a larger, upper level ontology – the CIK ontology, where all the central concepts of the Construction area are structured. Therefore the solution proposed is to develop an inter-organizational KM system for CIK Communities which will be built upon distributed ontologies locally managed and centrally integrated, as described in the following figure.

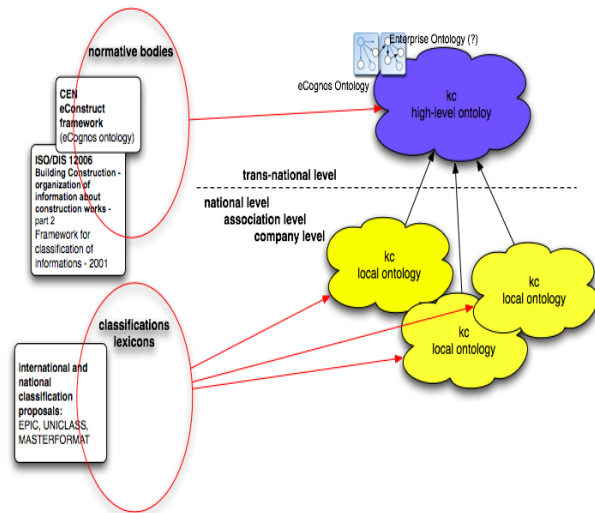


Fig. 2. General structure of the KC CIK ontology

#### 4.1 Managing multiple ontologies

CIK ontology integrates knowledge from several sources based on different conceptual models. The problem of semantic heterogeneity between these models must be solved. Today there are some established approaches to the process of managing distributed ontologies, usually named as integration.

According to Pinto *et al.* (1999), integration corresponds not only to the design of special operations to construct ontologies from others available in some ontology development environment, but also to the process of ontology construction based on other previously built, to the set of activities undertaken by some methodologies that specify new ontologies using publicly available ontologies, to the use of ontologies in applications, etc.

For the purpose of managing multiple ontologies a distinction has to be made between: (1) ontology merging, (2) ontology integration and (3) ontology use/reuse:

1. **Ontology merging** is the process of building a single ontology through the merging of several source ontologies. Usually, the source ontologies cover similar or overlapping domains. In this case, one wants to build an ontology merging ideas, concepts, distinctions, axioms, etc., that is knowledge from other existing ontologies on exactly the same subject.
2. **Ontology integration** is the process of creating a new ontology from two or more ontologies by overlapping the common parts. The domains of the source ontologies are different from the domain of the resulting ontology, but there is a relation between these domains.
3. **Ontology use/reuse** is the process of building an application using one or more ontologies. In this case, one wants to introduce into an application one or more ontologies that underlie and are shared among several software applications or one uses one or more ontologies to specify or implement a knowledge based system.

In the first two processes, merging and integration, the result will be a single ontology. In the using/reusing process, different ontologies are used but they are not unified into a single one. To help this process, there are techniques like mapping or aligning, which can be followed according to each situation.

#### 4.2 Managing KC ontologies

Taking into consideration the distinction made in the previous section, the ontology merging process was used to build the CIK ontology from the eCognos ontology, the ISO 12006 and EPIC that cover similar and/or overlapping domains in the Construction area. The intention was to build an ontology that merges the existing ideas, concepts, vocabulary, projects, etc., of the domain. CIK ontology is also the result from the agreement between all partners involved in the development of KC project and is to be used as an high-level ontology by the Associations in the semantic portal implemented by the KC system.

From the requirements analysis it was concluded that different countries (or even different associations) need to use local classifications, thesauri or ontologies. As different countries have different local ontologies, the best process to “plug-in” a local ontology into the CIK ontology is to align both ontologies. This will have the advantage of being easier to maintain and evolve the ontologies. In practical terms the option for ontologies integration can be a lot more difficult when managing the ontologies and its evolution process, for example to discover the part of the integrated ontology that will suffer changes.

To align CIK and local ontologies it is necessary to list the differences between their terms, understand their similarities, to detect possible inconsistencies and/or missing of information. This technique is particularly concerned with concepts with the same meaning, but labeled with different names or vice-versa; differences in the spelling of the terms, i.e., one term in plural and another in singular, one in the feminine and another in the masculine, or different verbal terms; properties with the same meaning, but labeled with different names or vice-versa; differences in constraints and properties used; etc.

The ontology alignment process is done semi-automatically (the human intervention is needed to make some decisions in different steps) using the open source PROMPT tool. PROMPT is used as a plug-in for Protégé editor, as all ontologies are being developed in OWL language using this editor.

#### 4.3 Developing CIK ontology

The central ontology (CIK ontology) reflects standards and related classification schemes in the industry and the local ontologies will account for the

individualised SME conceptual schemes, i.e. they will be strongly related to the consortium partners' needs.

The following description respects only to the CIK Ontology that defines the domain of the CI which can be summarized in the following sentence: *The Construction Industry involves a set of resources (Construction Resource) that follow certain conditions (Technical Topic) which are used or required in a process (Construction Process) that leads to results (Construction Result)*. As such, the proposed taxonomy includes four major domains to classify these major concepts: Construction Resource; Construction Process; Construction Result; Technical Topic.

As it is perceptible the first three domains coincide with the major themes in the ISO 12006-2 standard. The other domain (Technical Topic) is the result of the integration of an e-COGNOS module, further developed so as to include issues related to the CI that are not covered by the e-COGNOS ontology and IFC model. All domains integrate modules from eCognos ontology.

## 5. CONCLUSION AND FUTURE WORK

Given the selection, evaluation and structuring process described in this paper, there is a high level of probability that the CIK ontology reflects the standards and related classification schemes in the industry on the one hand and, on the other, the local more specific ontologies will account for the individualized SME conceptual schemes, i.e. they will strongly relate to consortium partners' needs, as identified in the analysis of business case scenarios and in the users' requirement definitions.

The implemented method aims at developing a methodology of common Construction Industry Knowledge representation applicable to large sets of SMEs in the construction industry as a basis for the establishment of a knowledge community. Because of the available time frame, the described method was not as fine grained as desired. Therefore, further work in this area will be directed to detail the evaluation criteria. The next steps will involve the definition of the ontologies and maintenance strategies. In parallel, the first prototype validation by the users will be setup.

## REFERENCES

- CWA/CEN Workshop Agreement, *European eConstruction Ontology (EeO)*, (2004). Documents produced by the CEN/ISSS eConstruction Workshop, Brussels.
- De Vries, S. and P. Kommers, (2004). "Online knowledge communities: future trends and research issues." *International Journal of Web Based Communities* (1): 115 - 123.
- Doan, A. et al., (2002) "Learning to Map between Ontologies on the Semantic Web," *Proc. World-Wide Web Conf.*, ACM Press, pp. 662–673. *Proc. 1st Int'l Conf. Computational Logic* (CL 2000), LNCS 1861, Springer-Verlag,
- e-COGNOS Public Deliverables, available at <http://www.e-cognos.org/>
- eConstruct Public Deliverables, available at <http://www.bcxml.org/>
- ISO 12006-2 Building construction — Organization of information about construction works - Framework for classification of information, DIS version 2001.
- Hearn, P., A. Bradier, et al., (2002). "Building communities: organizational knowledge management within the european comission's information society technologies programme." *ITcon 7*: 63-68.
- FUNSIEC, Feasibility study for an UNified Semantic Infrastructure in the European Construction sector – Public documents available at [www.funsiec.org](http://www.funsiec.org)
- Gómez-Gauchía, H., B. Díaz-Agudo, et al., (2004). Two-layered approach to knowledge representation using conceptual maps and description logics. First Int. Conference on Concept Mapping, Pamplona, Spain.
- Gronau, N. and E.-M. Kern (2004). Collaborative Engineering Communities in Shipbuilding. Virtual Enterprises and Collaborative Networks (PROVE'04), Toulouse, France, Kluwer Academic Publishers.
- Lima, C., (2004). Framework to Support Interoperability among Semantic Resources.
- Lima, C., Fiès, B., Ferreira-da-Silva, C, (2004). "Setting up the Open Semantic Infrastructure for the Construction Sector in Europe – the FUNSIEC Project." In: *5th European Conference on Product and Process Modelling in the Building and Construction Industry – ECPPM 2004*, Istambul, Turkey.
- Maedche, A. et al., (2002). "MAFRA—Mapping Distributed Ontologies in the Semantic Web," *Proc. 13th European Conf. Knowledge Eng. and Management* (EKAW 2002), Springer-Verlag, pp. 235–250.
- Pinto, Helena S.; and Martins, João P., (2001); *Ontology Integration: How to perform the Process*. Proceedings of the international conference on Knowledge capture. Victoria, British Columbia, Canada.