SELECTING AND STRUCTURING SEMANTIC RESOURCES TO SUPPORT SMEs KNOWLEDGE COMMUNITIES

António Lucas Soares
INESC Porto, Rua Roberto Frias, Campus da FEUP, 4200 Porto Portugal
University of Porto, Faculty of Engineering, Rua Roberto Frias s/n, 4200 Porto Portugal

Manuel Moreira da Silva
INESC Porto, Rua Roberto Frias, Campus da FEUP, 4200 Porto Portugal
ISCAP - IPP, Rua Jaime Lopes Amorim, s/n, 4465 S. Mamede de Infesta Portugal

Dora Simões
INESC Porto, Rua Roberto Frias, Campus da FEUP, 4200 Porto Portugal
University of Aveiro, ISCAA, R. Associação Humanitária dos Bombeiros de Aveiro, 3811 Aveiro Portugal

Keywords: Semantic resources, Knowledge communities, Ontology engineering, Ontology reuse/integration, KMS.

Abstract: Knowledge management intrinsically involves communication and information sharing, which can be strongly affected by the context in which it is viewed and interpreted. This situation gets worst when complex domains are considered, as it is the case of the Construction Industry domains. 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 decided, as an approach to this task, to re-use, as far as possible, existing ontologies, classification systems and other semantic resources in order to develop a system that may come to contribute to standards and to 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. It aims at developing a methodology of common Construction Industry Knowledge (CIK) representation applicable to large sets of SMEs in the construction industry as a basis for the establishment of a knowledge community.

1 INTRODUCTION

Building and construction companies have to continuously renew their working habits in order to face an increasing competitive environment where flexibility and adaptability to change are the obliged route to success.

The main challenge is to provide a cost-effective solution for the two main problems:

1. Construction industry (particularly SMEs) urgently needs radical improvements of communication with customers in order to provide better product support and services. The innovative forms of communications and relationships among SMEs and their customers are increasingly important in order to improve the market share and/or survival chances in the "new economy era".
2. To respond to ever increasing customer requirements it is increasingly necessary to establish a closer co-operation (particularly among SMEs) within this sector, aiming at assembling alliances of SMEs into integrated teams that will genuinely align with challenging performance targets.

Lucas Soares A., Moreira da Silva M. and Simões D. (2006).
SELECTING AND STRUCTURING SEMANTIC RESOURCES TO SUPPORT SMES KNOWLEDGE COMMUNITIES.
In *Proceedings of the Eighth International Conference on Enterprise Information Systems - AIDSS*, pages 270-277
DOI: 10.5220/0002465602700277

Taking into account these needs the the Know-Construct (KC) EU project[1] aims to develop an Internet Platform for Knowledge-based Customer Needs Management and for Support to Knowledge Communities of SMEs in Construction Industry. It involves professional associations of the sector as they may provide an ideal environment to disseminate innovative solutions.

The goal of KNOW-CONSTRUCT (KC) (www.know-construct.com) is to provide a common internet-based platform for SMEs from the construction sector to achieve an effective combination of two general functionalities:

1. *Customer Needs Management (CNM) System*: a decision making support system regarding the products characteristics, applications and other consultancy services for SMEs customers applying the "web enabled dialogue", and
2. *Knowledge Communities Support (KCS) System*: a system for SMEs to support a form of co-operation through the creation of Knowledge Communities of SMEs in Construction Industry.

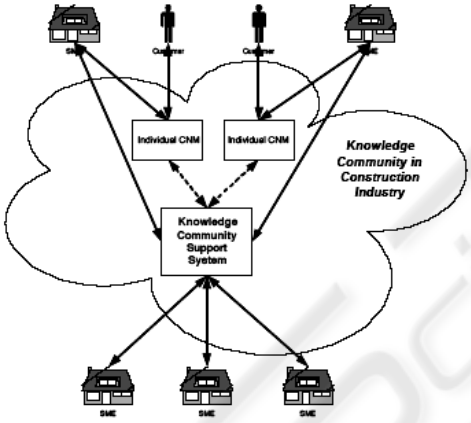

Figure 1: Know-Construct basic concept.

The system supports the integration, management and reuse of the area specific knowledge via a common knowledge base. The platform will be owned by associations, offering to members (SME) an area to establish specific customer relationship management (using the KC individual CNM systems) that take benefit of a knowledge community supported by the KCS System.

For a successful solution it was necessary to establish a multidisciplinary approach combining research activities in different areas, such as knowledge representation, particularly creation of the adequate ontologies and their structure, effective forms of on-line interaction among customers and SME (eConsultancy - web enabled dialogue with expert/technical advisor), as well as product classification systems for this sector, new forms for a representation of the CI knowledge (e.g. experience-based), investigation of efficient approaches for training, etc.

This paper describes a fundamental phase of the project which was the selection and structuring of the semantic resources (ontologies, classifications, thesauri), that form the cornerstone of the KC system. The rest of the paper is as following: section 2 describes briefly in the concept and functional architecture of the KC system; section 3 describes the process of selecting the relevant knowledge sources for the KC ontology construction and integration; section 4 describes the KC ontology structure and gives examples of its content.

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 of them. This area has not been sufficiently mapped before, but [Cornejo 2003] explores it very well 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 derive the maximum value to individual SMEs from a knowledge community where those companies participate.

2.1 Models for Knowledge Communities

KC general objective is to improve the relationship of CI SME's with their customers by providing 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 SME's. This knowledge community (Construction Industry

[1]EU - COLL-CT-2004-500276

Knowledge (CIK) community) can be considered as a hybrid of community of practice and interest, which is Organization-sponsored (Porter, 2004). From one side, company employees as individuals should see a direct utility to their particular jobs when participating in the CIK community. This direct utility comes into light when an employee (and consequently the company) realizes that, when solving a problem to 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 to a causal benefit to the SME. For example, some chatting between two employees exchanging a specific professional experiences or a report in a news or 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 of a concrete immediate value for the organization.

Based on work of de Vries *et al.* 2004, we synthesize the following characteristics of the CIK Community:

- The goal is to **develop and exploit knowledge** about construction area.
- There are **continuous interactions 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.

Following this explanation, CIK Community will consist 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** are all the professionals of the construction area and belonging to the associations of the project partnership.
- **On-line Professional Meeting Place** is the online place where the members of the community will take theirs discussions, debates, conversations, etc.
- **Organizational context** is related with the organizational goals, culture, technology, etc.

2.2 KCS System General Architecture

As mentioned before, the KC project has a very specific goal: to enable individual SME's to better 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, but focused fundamentally in generating broader and deeper knowledge to be used in managing the SME's 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, it 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

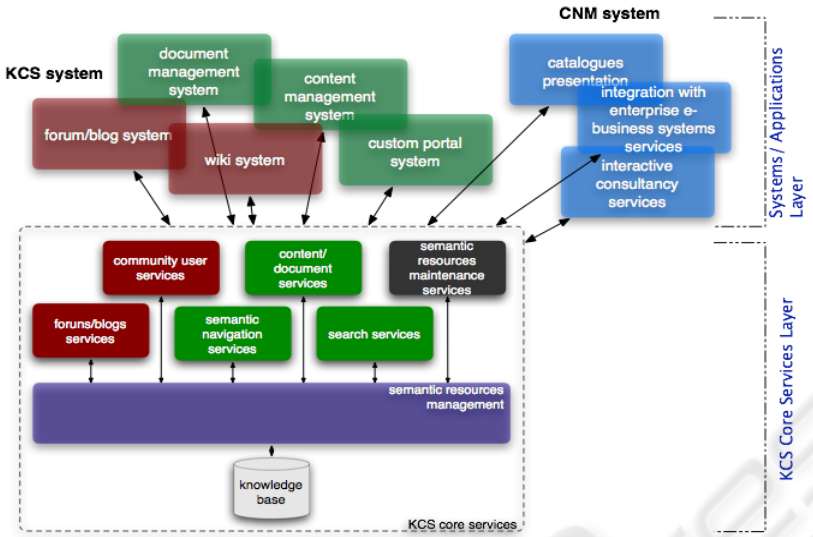

Figure 2: The Know-Construct functional architecture.

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 information/knowledge will be created in communication/interaction processes (foruns, weblogs), there will be also the need to create/access knowledge in a more structured way. Digital content management and document management are the natural approaches regarding this issue.

This generic architecture is represented in more detail in Figure 2.

3 SELECTION OF KNOWLEDGE SOURCES FOR KCS IN THE CONSTRUCTION INDUSTRY SECTOR

The identification and selection of existing knowledge sources was the first step to build the semantic resources structure of KCS system. Multiple different sources' categories 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 and design principles.

3.1 Types and Characteristics of Knowledge Sources

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 widely 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 our 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's 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 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, 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* - Methodology, tools and architectures for electronic consistent knowledge management across projects and between enterprises in the construction domain; *e-Construct ontology* - Electronic Communication in the Building and Construction Industry; *EPIC* – European Product Information Co-Operation; *UNICLASS* – Unified Classification for the Construction Industry; *IFC Model* – Industry Foundation Classes; *ICIS LexiCon* – International Construction Information Society.

The identification of these sources led 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 adapt the selected resources according to 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); the previously defined scenarios.

Table 1: Criteria for the evaluation of knowledge sources to be used in KC.

Type of source	***Common Criteria***	***Description***
All	Origin	Developer(s); type of entity (CEN, ISO, DIN, other)
	Relevance	for the pre-defined areas of analysis; for specific cultural and professional
	Adequacy	from the domain expert point of view; from the ontologist point of view
	Completeness	explicit in-depth coverage
	Comprehensiveness	domains addressed in the area
	Ease of data	Possibility of access and reuse (merge/integrate)
	Language	language(s) in which it is available, multilingual features, language
	Current status	finished, work in progress, in revision
	Specific Criteria	
Ontology	Conceptual	ontology assumptions and ontological commitment and their relation to KC
	Type of concept	identification of generic concepts and relationships, identification of 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	type of documentation available and accessibility
	Consistency	consistency of the application of the relations
	Modularity	which concepts are represented in which modules
Terminology	Terminology purpose	operational terms – functions the terminology is intended to serve
	Standardized/non-	implemented as standard, other type
	Granularity	level of complexity of the available data
	Quality of the	do they follow unified patterns, are simple, clear, concise, etc.
	Interconnectivity	to what extent is the terminology mappable to coding systems or
	Precision and recall	retrieval effectiveness
	Normalization	of content and semantics
	Responsiveness	Frequency of update
Classification	Classification	classification purposes and their relation to KC objectives
	Conceptual framework	classification assumptions and their relation to KC objectives
	Classification scope	Domain(s)
	Type of concepts	degree of abstraction/specificity
	Previous use	use in ontology projects and outcome analysis

3.2 Evaluation and Selection of Knowledge Sources

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 analysed knowledge sources in what concerns to their suitability to the KC system and methodology.

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

- elements and indicators previously identified,
- 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 process of identification, classification and evaluation, and taking in consideration the general view expressed in the CWA 15142 – European eConstruction Ontology (EeO), from the available 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: LexiCon, bcBuildingDefinitions, eCognos, IFC model, ISO 12006 - 2 and 3 - *Building construction — Organization of information about construction works*

Table 1 describes the criteria established and their scope inspired in the works of Pinto and Martins (2001) and Lelkin (2004).

4 ONTOLOGIES STRUCTURE AND CONTENT

After the activities described above, we delineated the adequate strategy in order to support the structuring, maintenance and evolution of CIK ontology and local ontologies. This level of ontology management is necessary not only for the initial development and maintenance of ontologies, but it is essential during deployment, when scalability, availability, reliability and performance are absolutely critical.

As it had been defined in the KC proposal, the methodology to be followed aims mainly at identifying, evaluating and reusing existing semantic resources, like ontologies in the IC area, due to the advantages their reuse offers: simple design, more reliable knowledge sharing and clearer semantic representations. New developments made in the construction industry context or in any other will be submitted to the above listed representative type of ontologies.

4.1 The Local/Global Semantic Resources Approach

During the analysis described above, a perception gained strength – to deal with the very concrete reality of the Associations and SMEs of each country, KC system would need to have a local ontology that would answer the KC partners' particular professional and cultural needs.

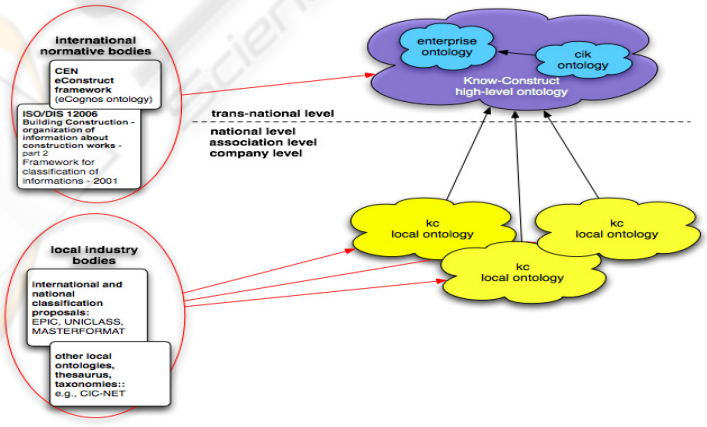

Figure 3: General structure of the KNOW-CONSTRUCT CIK Ontology.

The development of this more specific ontology will be based on a larger, upper level ontology – the CIK ontology, where all the central concepts of the Construction Industry area are structured.

The perception of the need to provide tailored support to the users was reinforced by the conclusions of the CWA 15142 on European eConstruction Ontology (EeO), where it is clearly stated that the e-COGNOS vision over the development of a big ontology was confronted with an unexpected reality. The end users actually showed their preferences to use their very specific, concise and precise taxonomies. They did not want to handle big ontologies; rather they are perfectly happy if their small resources are in place providing the results they are expecting. This fact has changed the concept of the e-COGNOS ontology: the big ontology is in place, but it is totally customisable in the sense that a small taxonomy with 100 concepts can replace the big one.

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 an away as to keep a common central ontological content (structure, attributes, relations, etc.) from where to derive the more specific ontologies.

Therefore the solution proposed is to develop an inter-organizational KM system for Construction Industry Knowledge Communities which will be built upo8n distributed ontologies locally managed and centrally integrated.

The central ontology reflect 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. This methodology results in a need to develop the two types of ontologies in two different moments which leads to two main concerns: 1) how to establish through the re-use and integration of existing ontologies (as far as possible) an adequate domain related ontology, as well as classification system for this sector applicable in SMEs environment, and 2) how to assure the continuous update/maintenance of both types of ontologies in order to enable a long life to the knowledge systems.

4.2 Building and Integrating the Global Ontology

Ontologies represent shared knowledge between the parties, and result from a shared approach to a knowledge domain. In the case of KC, the decision to re-use and integrate available ontologies and classifications standards in the CI area, led to an even greater need to develop a Know-Construct high-level ontology that would allow the integration of these resources. In order to better answer the needs and purposes of the CNM and KCS systems, this ontology (named as CIK Ontology) is integrated in the Enterprise Ontology, as defined by the Enterprise Project by the Artificial Intelligence Applications Institute at the University of Edinburgh.

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 can be seen 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.

4.3 Using and Integrating Local Semantic Resources

Other well developed resources that may be useful for the CIK ontology definition and further development are the country specific semantic resources. These resources are highly developed and have been used in previous projects in the CI area and, in the majority of the cases; they are the result of the work of several combined institutions and actors interested in developing the existing standards and classifications.

These sources are mandatory for future developments of the local ontologies, for the specific markets and specific Costumer Needs Management and Knowledge Community Systems, the sources have to be defined according to the local market context, SMEs and the Industrial Association/Grouping needs.

They need to be an integrant part of the local market information resources, the SMEs internal documentation and catalogues, internal documents and databases (contacts, material properties, specifications, standards, prices), and the diverse

publications in the area of the CI, especially those edited by the Associations or available in the internet (portals, individual sites, online libraries, newsgroups, etc.).

On a second stage of the ontology definition and management – the establishment of the local more specific ontologies - other more detailed local specifications/structures will have to be considered and developed taking into account the local/national sources of knowledge and of knowledge dissemination, also as a way to better address the management of tacit knowledge and, in particular, the social aspects of knowledge exchange.

As these sources are based on different conceptual models, in order to integrate them we must solve the problem of semantic heterogeneity between them. From the review on different approaches to support this task, we've decided to follow the ontology mapping approach that allows relating a portion of the source ontology to the target ontology's entities, transforming instances from the source ontology into instances in the target ontology. Ontology mapping approach allows transforming information but doesn't require the building of an integrated view. So, although it's more powerful than simple inclusion, it avoids the complexity and overhead of integrating multiple sources.

5 CONCLUSIONS AND FUTURE WORK

Given the selection, evaluation and structuring process described in this paper, there is an 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 be strongly related 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 for 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 we would like it to be. Therefore, further work in this area will be directed to detail the evaluation criteria.

Next steps will involve the definition of the ontologies and classifications maintenance strategies. In parallel, the first prototype validation by the users will be setup.

REFERENCES

CEN - European Committee for Standardization; CWA 14946:2004 E, European eConstruction Framework (EeF) - Context and Scope for "eConstruction", Brussels, 2004.

CEN 2003, CEN/ISSS Roadmap for addressing key eBusiness standards issues 2003-2005, source: CEN/ISSS eBusiness Standards Focus Group. http://www.eeuropestandards.org/Docs/

Cornejo, M. 2003. Utility, value and Knowledge Communities. Knowledge Communities (to be published).

CWA3 – CEN Workshop Agreement, *European eConstruction Ontology (EeO)*, 2004. Documents produced by the CEN/ISSS eConstruction Workshop, Brussels, 2004.

De Vries, S. and P. Kommers 2004. "Online knowledge communities: future trends and research issues." International Journal of Web Based Communities

e-COGNOS Public Deliverables, available at http://www.e-cognos.org/

eConstruct Public Deliverables, available at http://www.bcxml.org/

FUNSIEC, Feasibility study for an UNified Semantic Infrastructure in the European Construction sector – Public documents available at www.funsiec.org

Gronau, N. and E.-M. Kern 2004. Collaborative Engineering Communities in Shipbuilding. Virtual Enterprises and Collaborative Networks (PROVE'O4), Toulouse, France, Kluwer Academic Publishers.

ISO 12006-2 Building construction — Organization of information about construction works - Framework for classification of information, DIS version 2001.

Lima, C., A Framework to Support Interoperability among Semantic Resources, 2004.

Maedche, A., B. Motik, et al. 2003. "Ontologies for enterprise knowledge management." IEEE Intelligent Systems (Intelligent Information Processing): 2-9.

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.