Semantic Document Indexing in Ontology-Driven Organizational Memories

Mariel Alejandra Ale[1], Omar Chiotti[1,2], Maria Rosa Galli[1,2]

[1] CIDISI – GIDSATD, Universidad Tecnológica Nacional FRSF,
Lavaise 610, Santa Fe, Argentina
+54 0342 4608585
male@frsf.utn.edu.ar
[2] INGAR - CONICET,
Avellaneda 3657, Santa Fe, Argentina
+54 0342 4534451
{chiotti, mrgalli}@ceride.gov.ar

Abstract. Effective document retrieval using domain knowledge and semantics is one of the major challenges in Information Retrieval. Over the last years, there has been a growing interest in ontologies as an artifact for human knowledge representation and a critical component in Knowledge Management, Semantic Web, and Business-to-Business applications. We have found that it is not easy to represent certain types of knowledge (skills or procedures) or to transform certain types of knowledge representation (knowledge contained in diagrams) into an appropriate ontological format. To overcome this problem, our proposal is to connect knowledge sources to the domain ontology associated with an Organizational Memory without forcing any transformation in the structure of the source itself. This connection will allow the semantic classification of knowledge sources so that when a user performs a query it is possible to recover the documents that have a higher probability of containing the answer.

Keywords: Domain Ontology, Organizational Memory, Information Retrieval, Semantic Indexing, Knowledge Management.

Workshop: II Workshop de Ingeniería de Software y Bases de Datos (WISBD)

1 Introduction

In the last decades, Knowledge Management has become one of the most critical factors for success in organizations. Shorter products life-cycle, globalization, and strategic alliances between companies demand a deeper and more systematic organizational knowledge management. Knowledge Management solutions based on Information Technologies are focused, in general, on Organizational Memories that comprise heterogeneous knowledge sources (structured, semi-structured, and informal) to facilitate access, distribution, and reuse of knowledge [1].

Organizational Memories are the means by which knowledge from the past is brought to bear on present activities resulting in higher level of organizational effectiveness [9]. An Organizational Memory comprises a variety of information sources where information elements of different kinds, structures, contents and media types are available [1]. An Organizational Memory should be able to control and access this heterogeneous information sources according to the user`s information needs. The utilization of a single Organizational Memory for the entire organization presents serious difficulties related to creation and maintenance processes of a single knowledge source. One possible solution is to associate every organizational knowledge domain with its own Organizational Memory and add an interface that enables knowledge retrieval from other domain Organizational Memory if it is necessary. In this particular type of Organizational Memory, the characteristics, attributes, and semantics of the Knowledge Objects, as well as the relationships among them are represented through a domain ontology [3].

Nevertheless, implementing an Organizational Memory requires solving two important problems: knowledge objects indexing and retrieval. The first will be the focus of this work and, in this context, knowledge must be modeled, structured, and related in an appropriate form so that it can be customized and integrated in a flexible way in users´ daily tasks. Nowadays, the problem is not the lack of information or knowledge, which is widely available in any organization, but the impossibility of manual tracking and analysis of innumerable and heterogonous knowledge sources. Different methods have been proposed in the literature to increase Information Retrieval systems effectiveness, especially methods to improve document representation, a key point in Information Retrieval. Several solutions regarding this problem include the selection of representative sets of weighted-terms and the development of richer data representations in order to get better query matching, especially in the area of Natural Language Processing.

An alternative to go beyond the “bag of words” strategy is to organize indexing terms into a more complex structure than bags. Depending on the degree of complexity we find different levels of knowledge representation. The simplest one is a taxonomy which is a subject-based classification that arranges the terms into a hierarchy. We can go further in complexity and create a thesauri, which basically take taxonomies as described above and extend them to make them better able to describe the world by not only allowing subjects to be arranged in a hierarchy, but also allowing other statements to be made about the subjects. Finally, the most advanced stage of knowledge representation correspond to ontologies that not only represent knowledge, but are also capable of defining new knowledge [11]. Ontologies have been proposed as the answer to modeling and structuring problems related to Knowledge Management. They provide a formal conceptualization of a particular domain that can be shared within and between organizations. They also provide a semantic basis for meaning definition and they are typically used to provide semantics for communication between people and machines. Therefore, they constitute the conceptual backbone of any sophisticated Knowledge Management infrastructure.

In the last years, there has been a growing interest in ontologies as an artifact for human knowledge representation and a critical component in Knowledge Management, Semantic Web, Business-to-Business applications, and many others application areas. Several research communities assume that ontologies are the appropriate modeling structure for knowledge representation. Nevertheless, little discussion has been made in relation to the range of knowledge that an ontology can represent successfully [6].

Clearly, it is not easy to represent certain types of knowledge (for example, skills or procedures) or to transform certain types of knowledge representation (for example, knowledge

contained in diagrams) into an appropriate ontological format. We have found that, in these cases, the possibility to efficiently retrieve and reuse this kind of knowledge is limited.

A common solution involving ontologies is to analyze documents in order to extract instances to populate the ontology. This population process generates a Knowledge Base that is used to answer user's queries. In our opinion, it is unlikely that a document containing, for example, a set of steps describing a certain task could be somehow transformed to be fed into the knowledge base associated to a domain ontology without losing some important knowledge. To overcome this problem, our proposal is to link knowledge sources to the domain ontology associated with the Organizational Memory without forcing any transformation in the structure of the source itself. Documents would be indexed by concepts that reflect their meaning rather than words considered as isolated tokens with all the ambiguity that they imply. This connection will allow the semantic classification of knowledge sources so that when a user performs a query it is possible to recover the documents that have a higher probability of containing the answer.

In the next section, we offer an introductory vision of domain ontologies and the problems associated with them. Next, we present our strategy for semantic document classification along with an example. Finally, we discuss conclusions and future work.

2 Domain Ontologies

Over the last decade, the word "ontology" has become a fashionable term in Knowledge Engineering community. Several definitions of ontology can be found and it can be seen that these definitions have changed and evolved over time [8][12]. One of the first definitions is the one given by Neches [16] who defined an ontology as the basic terms and relationships comprised in a vocabulary of certain area along with the rules to combine those terms and relationships to define extensions of the vocabulary. It is necessary to emphasize that according Neches´ definition an ontology not only includes the terms explicitly defined in it, but also the knowledge that can be inferred from it.

A few years later, Gruber [13] defined an ontology as an explicit specification of a conceptualization. This definition became the most popular in the ontological community literature and was slightly modified by Borst [5] who established that ontologies are formal specifications of a shared conceptualization. One of the most recent definitions states that an ontology provides the elementary knowledge and the needed infrastructure to integrate knowledge bases independently of a particular implementation [10].

Artificial Intelligence literature also contains several ontology definitions, many of which contradict one another [18]. In our case, we will define an ontology as an explicit and formal description of concepts in a particular domain, properties of each concept describing various features and attributes of the concept and restrictions for those properties.

It seems natural that corporate ontologies have a tendency to reflect the size of the organization that generates them. Therefore, they are likely to be large and partitioned according to the organization´s needs. Nowadays, ontologies are often related to specific domains that restrict their scope. In fact, this restriction is what gives ontologies a real usability: as long as ontologies are models, they are partial specifications of the world, and imply a normalization according to a given

point of view. In our case, the partition arises on the basis of internal knowledge domains that will probably but not necessarily reflect natural divisions within the organization such as Marketing, Human Resources, Research and Development, etc.

An important advantage provided by ontologies comes from the Information Retrieval area, where the availability of an ontology allows replacing the traditional keyword-based retrieval approaches by more sophisticated ontology-based retrieval mechanisms [14][21]. In fact, ontologies are often presented as silver bullets for the Semantic Web [10] and are expected to bring several benefits to Information Retrieval related to recall and precision, user assistance in query formulation, and retrieval from heterogeneous knowledge sources.

The remaining question is how well ontologies can represent all organizational knowledge. Alan Newell [17] characterizes knowledge as a behavioral phenomenon. He sees knowledge in terms of agent´s objectives, the actions that the agent might be capable of, and the means by which the agent makes a selection to achieve its goal. A key concept implied in this definition is that knowledge allows the agent to perform procedures in order to reach his goals, and that we can attribute knowledge to an agent because we observe him behaving in an apparent rational way in the real world.

This vision of knowledge goes beyond the notion of "specification of a conceptualization" of a concept enumeration and its relationships. From Newel's perspective, knowledge does more than just providing an enumeration of what exists in the world; it directly links objectives to actions. According to this, it is obvious that knowledge has an important procedural component. It is exactly at this point where a strict application of ontological representation shows its limitation.

An important part of organizational knowledge sources are documents containing a type of knowledge that cannot be efficiently represented with an ontology. This valuable knowledge, in the form of procedures, skills, diagrams, etc., would be lost if we only try to find ontological concept instances in it.

To sum up, most intelligent systems are designed primarily to respond to questions from a large corpus of knowledge. In these cases, ontologies provide most of the muscle necessary to construct complete systems. Nevertheless, in order to construct systems that solve problems according to this new knowledge perspective, conceptualization is not enough. We must link, from a semantic point of view, relevant organizational documents with a domain ontology that is the core of an Organizational Memory.

3 Semantic Document Indexing Strategy

As we said before, our goal is to link from a semantic point of view relevant organizational documents with a domain ontology that is the core of an Organizational Memory. We propose a strategy for semantic document indexing where ontologies are used as the main structure for the classification process. Our proposal relies in the hypothesis that domain ontologies contain all the relevant concepts and relationships in the domain. Although how ontologies are built up from the knowledge sources in the domain is out of the scope of this paper, it is necessary to explore the diverse ontology learning and population approaches since they will be the basis for concepts and

relationship identification. These extracted concepts and relationships will be the input to our semantic document indexing strategy.

There are several research works regarding concept learning from texts. The main exploited paradigms are contextual similarity [20] and the use of patterns that give some evidence of the presence of ontological concepts and relationships [15][7][19]. In relation to automatic annotation, there are some supervised machine learning approaches to automate information extraction [22].

However, in our opinion, these techniques rely on assumptions that are not completely compatible with our application scenario. In the first place, machine learning approaches that induce extraction rules for each concept from training data do not scale well for a large number of concepts. Secondly, to perform the annotation of a few hundreds of concepts, the needed training set is on the magnitude of thousands of examples, which constitutes an effort that few people are willing to make. Finally, machine learning based approaches assume that documents have a similar structure and content. This assumption is not realistic in our case given the heterogeneity of knowledge sources associated with Organizational Memories. Consequently, we believe that it is necessary to count on a new strategy for semantic document indexing that takes into account the inherent heterogeneity of knowledge sources. Moreover, we believe that semantic relationships and inferences can be used in order to make more meaningful document indexing.

In order to reach our goal we have developed a strategy that comprises several steps:

- Step 1: documents pre-processing to obtain a more standardized form of plain text.
- Step 2: application of different techniques to identify the presence of ontological concepts in the text.
- Step 3: ontological relationships identification.
- Step 4: semantic classification of the document.

Due to the inherent complexity associated with semantics, the complete automation of knowledge acquisition still remains in a distant future [23]. It is for that reason that our proposal presents a semi-automatic approach in which an ontology expert has to be a part of the process. To illustrate our strategy, we present an example using a Travel[1] ontology that contains concepts from the tourism area and an extract of a web page[2] of the same domain that is shown in Table 1:

Table 1. Web page extract

As a pearl of Bali, Kuta is one of the best beaches in Asia, with the only waves, which breaks over sand instead of a coral. This most popular tourist destination has succeeded in combining the need of local people and visitors. The original Kuta villagers have involved in the tourism industry for years. Kuta and surrounding offers various kind of accommodation, from simple and cheap accommodation provided by the locals to luxurious accommodation managed by international hotel chain. Kuta's seas are ideally best for surfing. Everyday both Balinese and foreigners are found along the golden sands of Kuta beach. Kuta area is also completed with various tourist facilities such as restaurants, pubs, bars, souvenir shops, tattoo parlors, travel offices, etc. Along the north of Kuta, Legian street offers a number of high quality boutiques, excellent restaurants, cafes and bars influenced by western style. Exclusive hotels can be easily found in Seminyak, further north of Kuta, while antiques wooden furniture shops in Jalan Raya Kerobokan.

[1] available at http://protege.stanford.edu/plugins/owl/owl-library/index.html
[2] http://www.baliforyou.com/bali/bali_beaches.htm

3.1 Text Pre-processing

It is generally accepted that not all words are equally significant for representing the semantics of a document because some of them carry more meaning than others. Usually, noun words are the ones that are most representative of a document content. Therefore, it is usually considered worthwhile to pre-process the text of the documents in the collection to determine the terms to be used as index terms [4]. In our case, the pre-processing module allows natural language text preparation. In this sense, we intend to obtain a plain text in which special characters, abbreviations and/or acronyms have been eliminated to prepare the text for subsequent steps. This substitution is based on a list of the most common special characters, abbreviations and acronyms that could be customized for a given domain. An example of this substitution can be seen in Table 2:

Table 2. Text pre-processing

Kuta area is also completed with various tourist facilities such as restaurants, pubs, bars, souvenir shops, tattoo parlors, travel offices, etc. /ETCETERA/

Other examples of possible substitutions of commonly found acronyms and abbreviations are presented in Table 3:

Table 3. Examples of commonly found acronyms and abbreviations

Acronym/Abbreviation	Substitution
R.O.	Room Only
B&B	Bed and Breakfast
P.P.	Per Person
St.	Street
H/B, F/B	Half Board, Full Board

This pre-processing will remarkably increase the efficiency of subsequent steps.

3.2 Tokenization and Lexical-Morphological Analysis for Concepts Identification

This step is divided into two main phases: on the one hand, the tokenization of the text and, on the other hand, the lexical-morphological analysis of each token. The tokenization consists of dividing the text into single lexical tokens. This is an important task that involves activities such as sentence boundary detection, simple white space identification, proper name recognition, among others. After tokenization, a lexical-morphological analysis has to be done using a POS (Part-of-Speech) tool. In our case, we use the POS tagger provided by GATE[3] (General Architecture for Text Engineering) which specifies if a term is a verb, an adjective, an adverb, or a noun. The GATE platform has been widely used as a basis for Information Extraction processes and content annotation management. It provides the fundamental text analysis technologies on which we have constructed our strategy.

[3] available at http://gate.ac.uk/

Since GATE´s POS tagger uses different tags from the ones used by WordNet[4] (a tool used in subsequent steps), it was necessary to define mapping rules between tags as is shown in Table 4:

Table 4. Mapping between GATE and WordNet tags

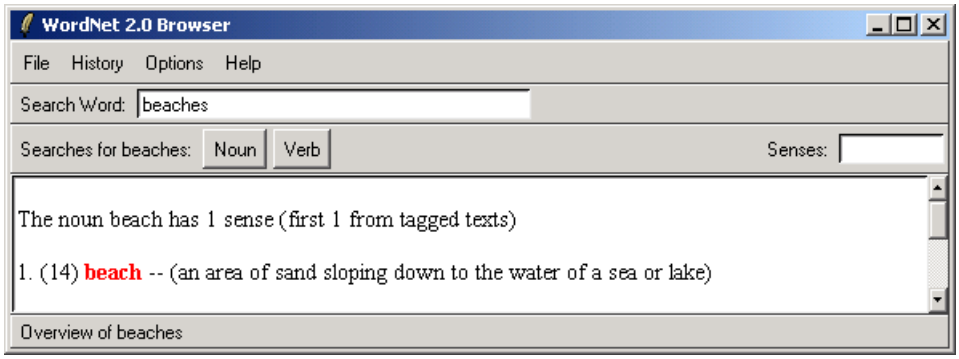
GATE tag	WordNet tag
NN: noun – singular or mass NNP: proper noun – singular NNPS: proper noun - plural NP: proper noun – singular NPS: proper noun - plural	<noun>
JJ: adjective JJR: adjective – comparative JJS: adjective - superlative JJSS: unknown – but probably a variant of JJS	<adj>
RB: adverb RBR: adverb – comparative RBS: adverb – superlative	<adv>
VB: verb – base form VBD: verb – past tense VBG: verb – gerund or present VBN: verb – past participle VBP: verb – non 3rd person singular present VBZ: verb – 3rd person singular present	<verb>

Usually, the decision on whether a particular word will be used as index term is related to the syntactic nature of the word. In fact, noun words frequently carry more semantics than adjectives, adverbs, and verbs [4]. As, in our case, index terms will be determined by ontological concepts, which are nouns, we will focus in this syntactic category within the tagged text. In this sense, ontological concepts can be seen as possible indexing categories. At this stage, using the synonyms set and hyperonymic/hyponymic structure provided by WordNet, we semantically expand every noun identified in the text and perform a search in the ontology. By doing this, we do not only identify exact ontological concepts occurrences but also derivations of the same word or even a synonym. Up to this point, we are not interesting in the meaning of each possible concept and that is why the presence of more than one sense for each noun in WordNet is not a problem.

In Figure 1, we present an example where the concept “beach” has been found with WordNet assistance. In this particular case, WordNet offers us the singular form of the word that constitutes an ontological concept. In other cases, this tool helps us to mark as ontological concept occurrences the presence of synonyms or other terms related with a “kind-of” relationship. In a way, if the noun is not found directly in the ontology, WordNet allows us to expand the matching possibilities taking advantage of related concepts (synonyms, hypernyms, hyponyms, etc.).

[4] available at http://wordnet.princeton.edu/index.shtml

Tagged Text:
"As a pearl<noun> of Bali<noun>, Kuta<noun> is one of the best beaches<noun> in Asia<noun>, with the only waves<noun>, which breaks over sand<noun> instead of a coral<noun>."



Ontological Concept Identified ====> beach

Fig. 1. Noun expansion for concept identification

Following the same methodology we found other ontological concepts in the original text: Destination, Accommodation, Hotel, Surfing and Beach.

3.3 Relationship Identification

In this step, as shown in Figure 2, we navigate through the domain ontology using the properties structure in order to find relationships among previously identified concepts. By doing this, we aim to contextualize those concepts that, in another way, could not be related with other concept among the ones that were identified in step 2. In the example shown in Figure 2, we retrieve two more concepts that do not belong to the original text – Activity and Sports – but provide a context for the "Surfing" concept.

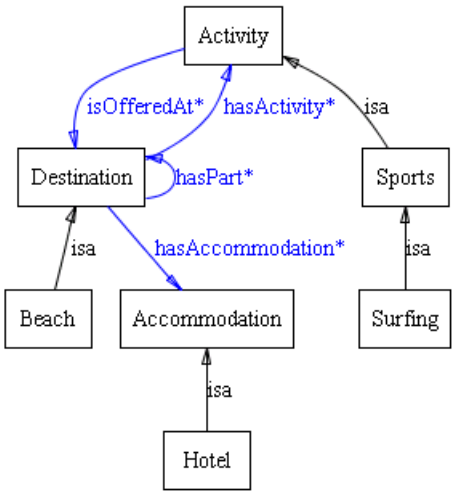


Fig. 2. Ontological document representation

We take advantage of the ontological relationships in order to perform a more accurate and contextualize representation of the document. As a result, we finally obtain the subset of the domain ontology that better models the document semantic content.

3.4 Semantic Indexing

Once all relevant ontological concepts and relationships are identified we link the document to the domain ontology by adding the document identifier as an instance of each related concept, as is shown in Figure 3:

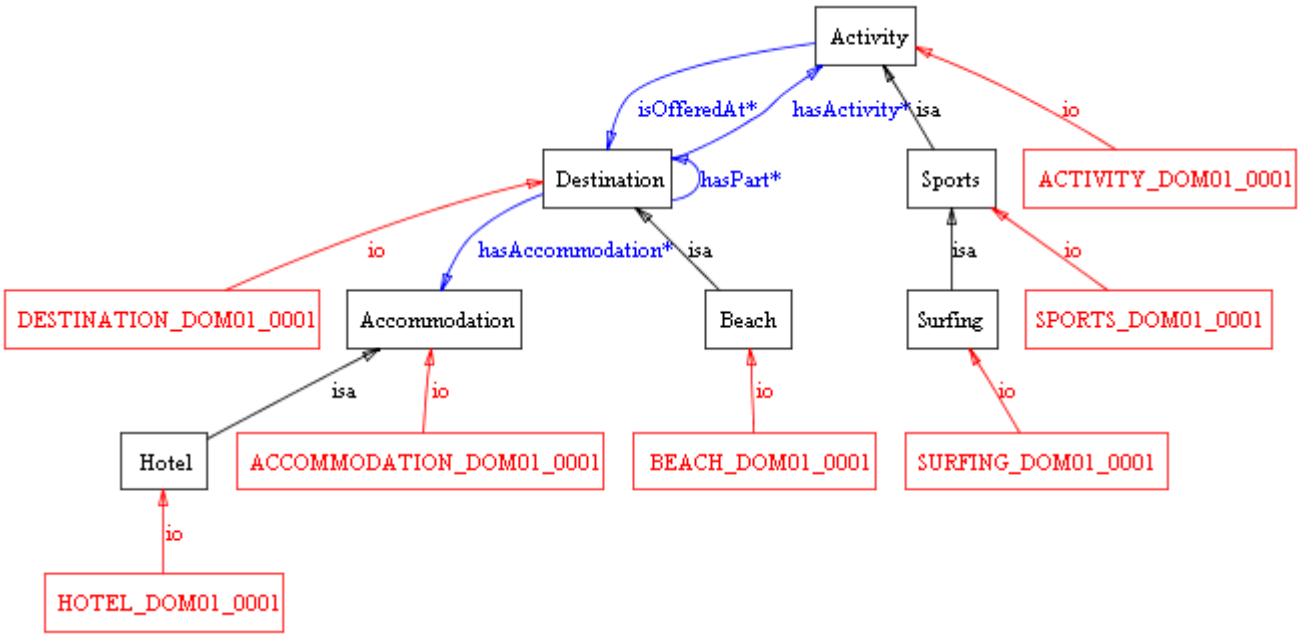


Fig. 3. Document linkage

Each instance identifier will be composed by the name of the instantiated concept, a domain identifier (due to the existence of various knowledge domains inside the organization, each one with its own domain ontology) and a unique document identifier. These identifiers will allow a subsequent ontology-based semantic retrieval of organizational documents. The advantage of this classification strategy based on domain ontologies becomes evident if we compare it with an approach based on keywords frequency. If we consider as representative those words with higher frequency, once we have eliminated stop-words, we obtain the words shown in Table 5 as key concepts:

Table 5. Keywords based on frequency

Frequency	Concept
8	Kuta
3	Accommodation
2	Tourist, Offers, Various, Found, Restaurants, Bars, Shops, North
1	Rest of the document

As it is shown, several key concepts that appear as relevant in the domain model constituted by the domain ontology are lost due to their low frequency. In addition, with this kind of statistical analysis it is not possible to establish relationships among identified keywords. Moreover, the ontology subset that models the document offers us much more information about the document content than a list of frequency-based keywords.

On the contrary, semantic document indexing enables new, semantically enhanced, access methods. In this way, the user could specify queries, which consists of constraints, regarding the type of concepts, relations between concepts and concepts attributes defined in the domain ontology. Once the query is processed, the system retrieves the corresponding documents and displays their identifiers organized according to the query terms.

4 Conclusions and Future Work

Information Retrieval based on keywords as index terms is simple but relies on the idea that the semantics of the documents and the user query can be effectively expressed through sets of index words. At least, this is an optimistic simplification of the problem because a lot of the semantics in a document or user request is lost when we replace its text with a set of words. For this reason, it is not uncommon that documents retrieved on a keyword basis are frequently irrelevant.

On the other hand, although ontology modeling is extremely valuable in the Knowledge Management field, the inherent limitations presented by domain ontologies to represent certain types of knowledge cause a large part of organizational knowledge sources to be unable to be incorporated into an Organizational Memory. It is for this reason that we have proposed an indexing strategy for knowledge sources (documents), independently of their content, that allows the linkage of these sources to the domain ontology incorporated in the Organizational Memory.

This strategy may be also applied to various domains and could be useful on the Web where it would be possible to structure an arbitrary set of web pages according to a particular view of the world given by the domain ontology chosen. The Web is a promising field of application since we see as an immediate step the necessity to apply this strategy to a larger corpus of documents in order to obtain statistical data for comparison purposes.

In previous sections we have abstracted from the relationship between ontology evolution and the semantic indexing strategy. However, in any realistic application scenario, new documents that have to be classified will generate the need for new concepts and relationships. Terms evolve in their meaning, or take on new meanings as organizational knowledge evolves. It is clear that we will have to find solutions to problems regarding with the addition, change or elimination of ontological concepts. We see the document indexing strategy and the ontology evolution as a cyclic feedback process that has to be taken into account in any Knowledge Management initiative.

References

1. Abecker A., Bernardi A., Junker M., Sarodnik T. and Sintek M.; KnowMore Project: Knowledge management for learning organizations, April 1997-March 2000, http://www.dfki.uni-kl.de/frodo/knowmore.html.

2. Ale M., Chiotti O. and Galli M.R.; Agent-Supported Ontology-Based Knowledge Management System. *Proceedings of ASIS 2004 (Simposio Argentino de Sistemas de Información) and JAIIO 2004 (33° Jornadas Argentinas de Informática e Investigación Operativa)*, Argentina, (2004).

3. Ale M., Chiotti O. and Galli M.R.; A Distributed Knowledge Management Conceptual Model for Knowledge Organizations. *Proceedings of JIISIC 2004*; Madrid, Spain; ISBN 84-688-8957-1; Page(s) 117-129; (2004).

4. Baeza-Yates R. and Ribeiro-Neto B.; *Modern Information Retrieval*; Addison-Wesley, Wokingham, UK, (1999).

5. Borst W.; *Construction of Engineering Ontologies*, PhD Thesis, University of Tweenty, Enschede, NL – Centre for Telematica and Information Technology; (1997).

6. Brewster C. and O'Hara K.; Knowledge Representation with Ontologies: The Present and Future. *IEEE Intelligent Systems.* Volume 19, Issue 1, Page(s):72 – 81; (2004).

7. Cimiano P., Pivk A., Schmidt-Thieme L. and Staab S.; Learning Taxonomic Relations from Heterogeneous Evidence. *Proceedings of 16^{h} European Conference on Artificial Intelligence*. Valencia, Spain; (2004).

8. Corcho O., Fernández-López M. and Gómez-Pérez A.; Methodologies, tools and languages for building ontologies. Where is the meeting point?. *Data & Knowledge Management*. Volume 46, Issue 1, Pages 41-64; (2003).

9. Croasdell D., Jennex M., Yu Z. and Christianson T.; A Meta-Analysis of Methodologies for Research in Knowledge Management, Organizational Learning and Organizational Memory: Five Years at HICSS; *36th Annual Hawaii International Conference on System Sciences (HICSS'03) - Track 4*; January 06 – 09; pp 110; (2003).

10. Fensel D.; *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce*; Springer-Verlag; Berlin; (2000).

11. Garshol L.; Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all; http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html#N1207; (2004).

12. Gómez-Pérez A., Fernández-López M. and Corcho O.; *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Ed. Springer-Verlag London Limited; (2004).

13. Gruber T.; A translation approach to portable ontology specification. *Knowledge Acquisition*. Volume 5, Issue 2; Page(s) 199-220; (1993).

14. Guarino N., Masolo C. and Vetere G.; Ontoseek: Content-based access to the web. *IEEE Intelligent Systems*. Volume 14, Issue 3; Page(s) 70-80; (1999).

15. Navigli R., Velardi P., Cucchiarelli A. and Neri F.; Automatic Ontology Learning: Supporting a Per-Concept Evaluation by Domain Experts. *Proceedings of 16h European Conference on Artificial Intelligence*. Valencia, Spain; (2004).

16. Neches R., Fikes R., Finin T., Gruber T., Senator T. and Swartout W.; Enabling technology for knowledge sharing. *AI Magazine*. Volume 12, Issue 3, Page(s) 36-56; (1991).

17. Newell A.; The Knowledge Level. *Artificial Intelligence.* Volume 18, Issue 1; Page(s) 87-127; (1982).

18. Noy N. and McGuinness D.; *Ontology Development 101: A Guide to Creating Your First Ontology*; Stanford University; (2001); http://protege.stanford.edu/publications/ontology_development/ontology101.pdf

19. Ogata N. and Collier N.; Ontology Express: Statistical and Non-Monotonic Learning of Domain Ontologies from Text. *Proceedings of 16h European Conference on Artificial Intelligence*; Valencia, Spain; (2004).

20. Reinberger M. and Spyns P.; Discovering Knowledge in Texts for the Learning of DOGMA-inspired Ontologies. *Proceedings of 16h European Conference on Artificial Intelligence*. Valencia, Spain; (2004).

21. Richard-Benjamins V., Fensel D., Decker S. and Gómez-Pérez A.; (KA)2: building ontologies for the internet: a mid-term report. *International Journal of Human-Computer Studies*. Volume 51, Issue 3, Page(s) 687-712; (1999).

22. Sabou M.; Extracting Ontologies from Software Documentation: A Semi-Automatic Method and its Evaluation. *Proceedings of 16h European Conference on Artificial Intelligence*. Valencia, Spain; (2004).

23. Staab S. and Studer R.; *Handbook on Ontologies*. Ed. Springer-Verlag Berlín Heidelberg; (2004).