Knowledge Engineering Suite:
A Tool to Create Ontologies for Automatic
Knowledge Representation in Intelligent Systems

Tânia C. D. Bueno,[1] Hugo C. Hoeschl,[1] Andre Bortolon,[1] Eduardo S. Mattos,[1]
Cristina Santos,[1] Ricardo M. Barcia [2]

[1] Instituto de Governo Eletrônico, Inteligência Jurídica e Sistemas – IJURIS
Rua Lauro Linhares, 728, sala 105, Florianópolis BRASIL – CEP 88036-002
http://www.ijuris.org
tania@ijuris.org;{hugo, bortolon, mattos, cristina}@wbsa.com.br

[2] Virtual Institute of Advance Studies – VIAS
Florianópolis, BRASIL
rbarcia@uol.com.br

Abstract. The present work is focused on the systematization of a process of knowledge acquisition for its use in intelligent management systems. The result was the construction of a computational structure for use inside the institutions (Intranet) as well as outside them (Internet). This structure was called Knowledge Engineering Suite, an ontological engineering tool to support the construction of ontologies in a collaborative environment and was based on observations made at Semantic Web, UNL (Universal Networking Language) and WordNet. We use both a knowledge representation technique called DCKR to organize knowledge, and psychoanalytic studies, focused mainly on Lacan and his language theory to develop a methodology called Engineering of Mind to improve the synchronicity between knowledge engineers and specialists in a particular knowledge domain.

1   Introduction

The importance of the Knowledge Based Systems is in the fact that they provide the computer with some peculiar characteristics of human intelligence, such as the capacity to understand natural language and simulate reasoning in uncertainty conditions. Defining the relevant information to be inserted into a Knowledge Based Systems is the great problem in the development of intelligent systems, mainly because the process is basically experimental and depends greatly on the ability of the knowledge engineer. In particular, a great difficulty is related to the definition of the terminology used to nominate the concepts and the relations [1]. Besides the great number of methods to do the knowledge acquisition, we can't find one that deals with the understanding and learning of the people involved, both specialists and knowledge engineers.

More recently, the notion of an ontology is being so popular in fields such intelligent information integration, information retrieval on the Internet, and knowledge management. The reason is in part due to what they promise: a shared and common understanding of some domain that can be communicated across people and computers [2]. Different developments of a worldwide range have a reference in cooperative work as a WordNet, Semantic Web and UNL (Universal Networking Language) through the construction of ontologies using collaborative tools. The use of ontological engineering tools or metatools to support the knowledge engineering process allows the organization of a knowledge base established on the relationship between relevant expressions from a context. Ontologies, as a basis for automatic generation of knowledge acquisition tools, simplify the tool specification process by taking advantage of ontologies defined as part of the knowledge engineering process [3]. Nevertheless, experience shows that often the bottleneck of building sharable ontologies lies more in the social process than in the technology [4]. For this reason, we develop a methodology to the process of knowledge acquisition to allow the specialist and the knowledge engineer to work in synchronicity, in cooperative networked organizations. We call this methodology Engineering of Mind. This synchronization process initiates with the understanding of human intelligence, its unconscious manifestations and its relationship with the words, since, according to Lacan, every human investigation is linked irreversibly in the inner space created by language. In the present development, we create a tool to support the knowledge engineering process by assisting developers in the design and implementation of ontologies in a specific domain.

In earlier works, we use a methodology called DCKR (Dynamically Contextualized Knowledge Representation [5]. DCKR allows the construction of a knowledge base, improving the construction of the domain ontology, and the automatic representation of cases in knowledge-based systems, either in the legal area [6], or in knowledge management domain [7].

In the next section, the methodology for the knowledge synchronization is described. This methodology allowed an exceptional coherence among the semantic relations of what is called 'indicative expressions', mainly by the support of all this computational structure during the process. This allowed the knowledge engineer and the specialist to develop much more than the knowledge representation of the domain, but abilities as inherent conscience, discipline, persistence, and empathy.

2   The Knowledge Representation in Knowledge Based Systems

We use a special process to extract and represent the knowledge for the knowledge based systems. The main intention of this process is to allow an automatic process of text indexing, on the basis of a controlled vocabulary and a dictionary of normative terms, constructed through the relevance of the definite terms persuasively, called normative key-terms [8]. Due to the necessity to turn the acquisition process quicker, it was necessary to evolve the process, using IR techniques (Information Retrieval) to associate the relevance of the terms with the frequency of the words added to the controlled vocabulary and the dictionary of normative terms; this approach gave origin to a methodology of knowledge representation called DCKR - Dynamically Contextualized Knowledge Representation [9].   DCKR is a methodology of

representation of knowledge whose approach is centered in a dynamic process acquisition of the knowledge of texts, defined through elaboration of a controlled vocabulary and a dictionary of terms, associated to an analysis of frequency of the words and indicative expressions of the context.

UNL, Semantic Web and Wordnet

In the knowledge acquisition for elaboration of the knowledge base of intelligent systems we chose the use of methodologies that use web environments and cooperative development. Today, there are three great worldwide developments that use the Internet for the development of ontologies, the UNL, Semantic Web and Wordnet.

UNL (Universal Networking Language) [10] is a language for computers to share information through a net. It is meant for representing the natural, independent language of its language, so that computers process the text and represent it in different languages.

The WordNet [11] is a lexicon reference system whose design is inspired in psycholinguistic theories on the human lexical memory. The nouns, verbs, adjectives and adverbs of the English language are classified only, being organized in sets of synonyms, each one representing a lexical concept. The sets of synonyms are through relations different to each other.

The Semantic Web [12] is an extension of the current Web, in which the information has a very well defined meaning, allowing the computers to process the information contained in the pages and to understand it, executing operations that facilitate the work for the users.

The three initiatives are meant to facilitate the automatic processing of the information contained in documents, allowing the computers to execute more intelligent operations and to retrieve information in more efficient way.

3   The Knowledge Engineering Suite

The Knowledge Engineering Suite is an Ontological Engineering Tool for collaborative networked works on the Web. Built to facilitate the knowledge sharing between the knowledge engineering team and the specialist team. The Suite allows to build relationships between complex terms, considering its concepts in the specific domain of the application. These relations are based on AI techniques [13], theories of language, Semantic Web, WordNet, and UNL.

The creation of an infrastructure for acquisition of the knowledge for cooperative work on the Web is an efficient and effective tool for the acquisition of the knowledge in intelligent systems. Many different techniques of Knowledge Acquisition exist; the Knowledge Engineering Suite (see figure 1) is used with the DCKR methodology. Where tools as the Frequency Extractor, Semantic Extractor and the Knowledge Engineering Suite have been associated to this methodology to assist in the task.

Fig. 1. Editing Module - Ontology construction (insertion and consistency checking).

This application works with extractors of automatic standards in contribution with knowledge engineers and specialists in the approached domain as specifications found in methodology DCKR - Dynamically Contextualized Knowledge Representation, which consists of a dynamic process of analysis of the general context that involves a thematic focused. The Suite is an editor of ontologies structured in a form to allow a cooperative work on the Web between the team of knowledge engineering and the team of specialists.

This computational environment of shared access has two main objectives: organization and representation of the knowledge, and update of the Knowledge Base.

Basically, four modules compose it, they are:

1. Register. It allows the elaboration of a contextualized dictionary, for the selection of topics and sub-topics for the classification of the indicative expressions. In this environment the user defines the topic and sub-topic in which it will insert a new indicative expression. A domain can be categorized in innumerable topics and sub-topics;

2. Search. It informs about other terms already registered in the base, which have some phonetic similarity with the term typed. This tool allows the verification of possible typing errors , besides preventing the registration of the same term more than once. It is a search system based on similarity. It supplies the user with a list of similar indicative expressions present in the knowledge base in alphabetical order after consultation made by the user. It is used in the registers, in the edition and the administration module;
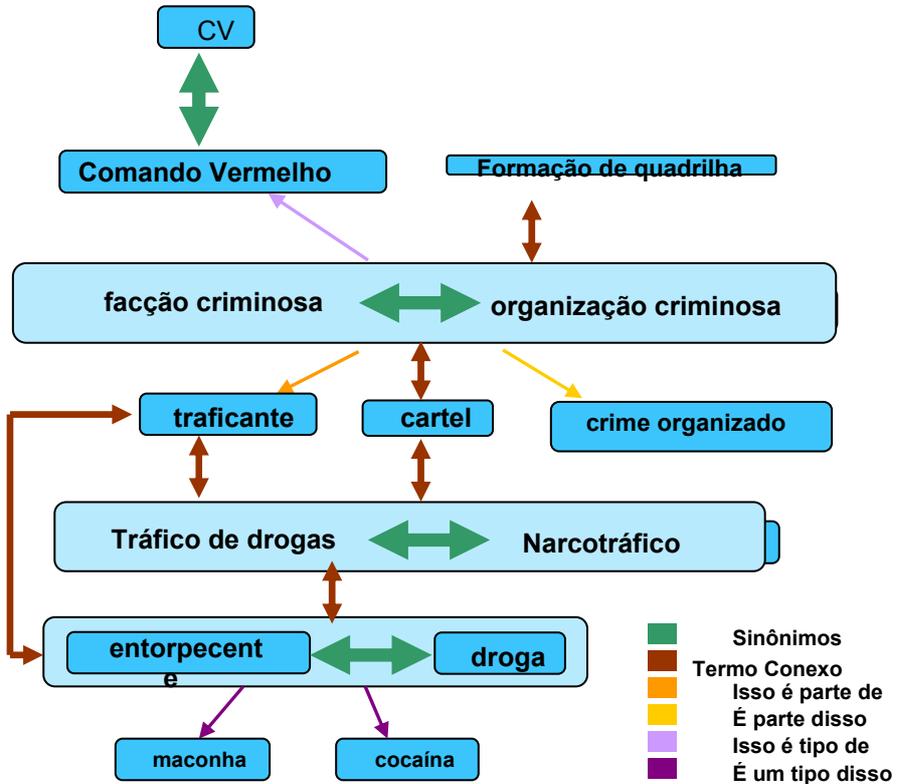
Fig. 2. The semantic relations of indicative expressions.

3. Relationship Editor. It allows the building of the relationship tree, always considering the similarity between all the terms registered and the ones already existing in the base. These relationships allow the system to expand the search context. The organization of the tree allows the dynamic definition of the weights of the indicative expressions according to the entrance of the user. The fields with all the relationships available to be formed are presented. They are the following: -*synonyms*; -*Related terms*; -*This is type of*; - *It is a type of this*; - *This is part of*; - *It is part of this*. The editor presents the existing relationships and allows excluding them (see figure 2). Each relationship has a weight related to the defined indicative expression in the search by the user.

4. Administration Environment. The knowledge integration and the validation between words is made in accordance with the context of themes and subthemes. This topic is organized in three levels: - High Level - it allows to insert themes and subthemes, to validate exclusions, to include and to exclude users, to verify productivity of each user and to verify descriptions of the dictionaries, themes, subthemes and indicative expressions; - Medium level- it allows to verify productivity and historical data; and, Low level- it allows to verify descriptions.

However, all this structure and methodology had not been enough to turn the cooperative work efficient and effective. It was necessary a more holistical approach,

which allows a greater coherence between the relations of the expressions, mainly in the elaboration of the related terms where the participation of the specialist is almost exclusive. It is important to highlight that this structure of contextualized ontologies allows an automatic information indexing by the system and an knowledge acquisition that gives more qualitative answers in the retrieval process.

4   Elaborating the synchronicity in a collaborative networked organization

The different unfolding of the human inventivity, even being so diversified, have the same origin, the unconscious mind and the human perceptions; from the fact of distinct constructions eventually to lead the thoughts to one same reference. Because of that, we elaborate a methodology that let the immediate perception of the specialist arise, without the pretension to reach all the knowledge, but with clear objectives, for example, to eliminate the common resistances of the people to the technological innovations, standing out the importance of the management of the human capital. [14].

During the development of tasks of knowledge engineering, it was observed that the efficiency of the acquisition process had a direct relation with good relationship between the knowledge engineer and the specialists of the domain, no matter the quality or content of the interviews, or the efficient application of the support tools. Thus, keeping this relationship in perfect synchrony is a key factor for the success of the system and a challenge for which the stages defined in the present work serve as a model of relative success.

The *Common-sense* tells us that the immediate perception (intuition) has greater effectiveness on the best solution for a problem than the application of rules of the propositional logic. Although, the most accepted proposal is people try to solve deductive problems applying rules such as of the propositional logic. According to Lacan [15], if we consider that the unconscious is structured as a language, it is possible to reconstruct the unconscious associations between the words, thus disclosing, a context.

There are elements, like the cognitive complexity and the capacity to learn, that supplies the underlying individual traces in which the specialized knowledge and abilities are based, similarly the sociability and the confidence supply the anchors to develop and to keep a net of relationships. Thus, identifying that non-cognitive knowledge is also important knowledge of the institutions and, for this reason, they must be part of the capital of these organizations, becomes necessary to look for a way to identify it and to represent it in the knowledge based systems. Therefore, this complex net of communication between the diverse areas of talent that will go to supply necessary flexibility, versatility and adaptability intelligences to happen.

All the languages are structurable as an articulating system. But their character, their coherence is in an articulated system, which is unique. Thus, the cognitive point of view concerns the symbolic acquisitions, those have as a foundation the meanings whose support is, generally, natural language, or, at times, specialized languages, as the formal ones. The attachment of these elementary meanings in a wide team requires synchronous thoughts.

This synchronization process initiates with the understanding of human intelligence, its unconscious manifestations and its relationship with the words, therefore, in accordance with Lacan [16], every human investigation is tied irreversibly in the interior of the space created by the language. But, for the victory of this dynamics of cerebral gymnastics, it is primordial that the person is in a positive attitude. The brain only registers, learns and ramifies when it is open to what is new.

4.1 Engineering of Mind Methodology

There are many different techniques of Knowledge Acquisition. We created the Engineering of Mind (see figure 3) to help developing the following process (DCKR methodology): 1. Inventory of the entire domain (classification of all sources of digital information that will be in the system database). 2. Application of the word frequency extractor based on the database inventoried; 3. Comparison between extractor results with the specialist's needs. 4. Construction of a representative vocabulary of the domain, by the specialist and knowledge engineers. 5. Application of the semantic extractor on the database; using the representative vocabulary (indicative expressions). 6. Definition of a list of words based on evaluation of the result of the frequency of the indicative expressions found in the inventory. 7. Construction of the ontologies in the Knowledge Engineering Suite based on this controlled vocabulary. 8. Definition of synonyms, related terms, homonyms, hyponyms, hypernyms and meronyms.

The acquisitions of the knowledge carried through by the team of engineers of the knowledge, in the area of its specialization [5] [6] got a bigger effectiveness than the acquisition carried through for the same team in diverse dominium of its specialization [7], where some obstacle of communication had taken the necessity of a new implantation of the acquisition process. That is, it did not have synchronization problems, therefore the deep knowledge of the specialists of the area of the technique of AI that was being applied in the system modeling (e.g., Case-Based Reasioning) allowed a transference of knowledge for the computational language of a very positive form for the final target of the systems.

It was observing the elements presented in the two processes that were possible to systemize a series of questions, improving the speed and quality of the knowledge represented in the system.

Associated to these comments, very uncommon procedures to the process of knowledge acquisition had been adopted, such as programming techniques neuro-linguistics and meditation to defragment the emotional memory of the specialist and to facilitate the learning process. This process was due to the existence of the following problems: 1. Resistance to the system; 2. Difficulty to reproduce the process of decision; 3. Little quality of the handled knowledge.

However, the focus object is not the area of application of the system (domain), but the specialist(s) and the knowledge engineer(s) that (will) exactly work in the definition of the target of the system and in the formation of the knowledge base of this system. To identify and to separate knowledge conditions are essential, therefore both (specialists and engineers) will have to learn and to train the learning process and, for this, it will need to surpass the comfort zone. Knowledge Engineering is over all knowledge exchange.
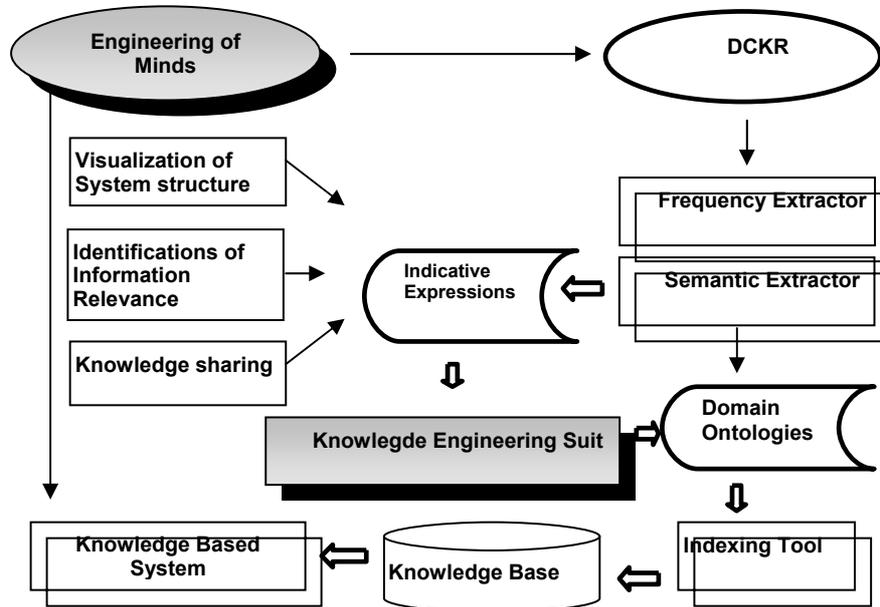
Fig. 3. Engineering of Minds Methodology applied to construction of ontologies
in Knowledge Based Systems.

The importance that the existing knowledge has for the new acquisitions comes from the basic role that they play inside of the construction of the representations and of the idea of that the acquisition passes, necessarily, by these representations. This is the importance of this phase, there is an exchange of knowledge, the specialist starts to know the form as its knowledge could be organized, that is, the basic concepts of the technique of used Artificial Intelligence in the representation of the knowledge, thus it will be able to contribute with more effectiveness and will have greater interest in participating in the process. As to the specialist, the exchange will lead to a more immediate perception of the target of the system, and will increase the interest in if going deep the study of the domain. Both will be prepared to deal with this overload and to obtain the ability necessary to plan or to choose a perspective that it determines then that elements of the situation must be treated as important and which can be ignored. When perceiving that of the vast information, the knowledge if only restricts to some of the number of characteristics and possibly excellent aspects, to take a decision one turns easier.

The continuous sharing of the established visions becomes the specialists and engineers better to work in cooperation in the construction of the ontologies of the domain. This productive process is continuous and can establish changes in elapsing of the implantation of the system.

5   Conclusions

The systematization and organization of teams of domain specialists together with the team of Knowledge Engineering started to be the great challenge in the system

development of knowledge management. The cooperative work between the teams does not only need the deep knowledge on the domain application to the system, but also on the organization of its knowledge base. The creation of a computational environment in the Web allowed to a greater sharing of information and better results among the teams in the construction of knowledge based systems.

   The Knowledge Engineering Suite enables a cooperative work among people in different places, structuring a continuous knowledge base and easy visualization (knowledge tree), through relationship nets and supplied an exceptional coherence among the semantic relations of what is called 'indicative expressions', mainly by the support of all this computational structure during the process. This allowed the knowledge engineer and the specialist to develop much more than the knowledge of the domain, but abilities as proper conscience, disciplines, persistence, and empathy.

References

1. Resende, Solange Oliveira. Sistemas Inteligentes: fundamentos e aplicações. Barueri, SP: Manole,2003.
2. Duineveld, A. J. et al, 1999. WonderTools? A comparative study of ontological engineering tools. *Twelfth Workshop on Knowledge Acquisition, Modeling and Management*.Voyager Inn, Banff, Alberta, Canada.
3. Eriksson, H. et al, 1999. Automatic Generation of Ontology Editors. *Twelfth Workshop on Knowledge Acquisition, Modeling and Management*.Voyager Inn, Banff, Alberta, Canada.
4. Benjamins, V.R., 1998. The ontological engineering initiative (KA)[2]*, Formal Ontology in Information systems*. IOS Press, Amsterdam.
5. Hoeschl, Hugo. C. Hoeschl, Bueno, Tania. C. D., Barcia, Ricardo. M., Bortolon, Andre., Mattos, Eduardo Da Silva. Olimpo: Contextual structured search you improve the representation council of UN security with information extraction methods In: å. Artificial International conference on intelligence and law, 2001, St. Louis. ICAIL 2001 Proceedings. New York: ACM SIGART, 2001, p.217 – 218.
6. Bueno, Tânia Cristina D'Agostini. O Uso da Teoria Jurídica para Recuperação em Amplas Bases de Textos Jurídicos. 1999. 94 f. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Florianópolis, 1999.
7. Ribeiro, Marcelo Stopanovski. KMAI, da RC²D à PCE. Gestão do conhecimento com inteligência artificial, da representação do conhecimento contextualizado dinamicamente à pesquisa contextual estruturada. [2004]. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2003.
8. Bueno, Tânia C. D. et al, 1999. JurisConsulto: Retrieval in Jurisprudencial Text Bases using Juridical Terminology. *Proceedings of the Seventh International Conference On Artificial Intelligence And Law*. ACM, New York.

9. Hoeschl, Hugo. C. et al, 2003. Structured Contextual Search For The Un Security Council. *Proceedings of the fifth International Conference On Enterprise Information Systems.* Anger, France, v.2. p.100 – 107.

10. UNL. Universal Networking Language. Available in: www.unl.ias.unu.edu/unlsys/index. html. Access in: 19 jan. 2004.

11. WORDNET. Available in: http://www.cogsci.princeton.edu/~wn/. Access in: 19 jan. 2004.

12. Semantic Web. Available in: http://www.w3.org/2001/sw/. Access in: 19 jan. 2004.

13. Kolodner, J. Case-Based Reasoning. Morgan Kaufmann, Los High, CA. 1993.

14. Gratton, Lynda, Ghoshal, Sumantra. Managing Personal Capital Human: new ethos will be the "Volunteer" Employee, The European Management Journal, vol 21, n° 1 pp1-10, February, 2003.

15. Lacan, Jacques. Os seminários de Lacan. Disponível em CD Room, 2000.

16. Miller Jacques-Alain, 1988. *Percurso de Lacan: uma introdução*. Jorge Zahar Editor Ltda, 2ª edição, Rio de Janeiro.