Knowledge Engineering for Non-Engineers

Tatiana Gavrilova

Knowledge Engineering for non-Engineers

Tatiana Gavrilova

Graduate School of Management - Saint Petersburg University
Saint Petersburg, Russian Federation

ABSTRACT. This paper presents one approach for the innovative organization training for business analysts in developing enterprise ontologies. The underlying teaching framework is pursuing a methodology that will aid the process of knowledge structuring and practical ontology design, with emphasis on visual techniques. The described approach may be helpful for those companies which are interested in the practical knowledge management and need skillfu knowledge workers. The paper proposes some new ideas of practical ontology design and evaluation and may be interesting for the knowledge engineering research and practicising community.

KEYWORDS: knowledge engineering, learning, thinking, analyst training, ontology design.

1. Introduction

During the last decade, knowledge has become a key consideration in our economies and it is heavily associated with learning and innovation. Central problems for supporting all phases of knowledge processing are the productivity of the knowledge workers and the effectiveness of the usage of the special professional techniques. These techniques and models help to elicit, stucture and integrate various knowledge patterns within and across enterprises. Knowledge work deals with analyzing and structuring in general. Top managers and IT analysts are continually challenged by the need to analyze massive volumes and varieties of multilingual and multimedia data. This situation is not limited to e-business, but is seen in nearly all companies and institutions. Knowledge base of a company can be operationalized, both in terms of measurement and by providing simulation models (Leydesdorff, 2006). Special interest to knowledge work is paid in the virtual and open oganizations.

Company staff and employees require support and guidelines for knowledge sharing about information analysis, theories, methodologies and tools. Knowledge management (KM)

is one of the powerful approaches to solve these problems in new information age with huge information overload and sophistication (Firestone and McElroy, 2005). Sophistication needs professionals. Professional knowledge analysts are still very rare on human resources market. Unfortunately, they also differ considerably in both backgrounds and cognitive styles (Wiig and Wiig, 1999).

Knowledge Engineering (KE) traditionally emphasizes and develops a range of techniques and tools including knowledge acquisition, conceptual structuring and representation models (Scott et al, 1994; Firestone, 2003). But for practitioners as enterprise analysts it is still a rather new, eclectic domain that draws upon areas like cognitive science. Accordingly, knowledge engineering has been, and still is, in danger from fragmentation, incoherence and superficiality. Still few universities deliver courses in practical knowledge engineering.

This paper describes recent experience in such training for some Russian subsidiaries of the international companies (British-American Tobacco, Siemens Business Services, etc.). The total number of trainees that received certificates of knowledge analysts is more than 60.

Theoretical part of the Training on Knowledge Engineering (TKE) is based on university courses in intelligent-systems development, cognitive sciences, user modeling and human-computer interaction delivered by author in 1992-2008 at the University of Pittsburgh (USA), University of Milano (Italy), University of Espoo EVTEK (Finland), Tartu University (Estonia), First Independent University of Warsaw (Poland) and Saint-Petersbutg State University (Russia). TKE proposes information structuring multi-disciplinary methodology, including the principles, practices, issues, methods, techniques involved with the knowledge elicitation, structuring and formalizing. Emphasis is put not on the technologies and tools, but in the training of analytical skills. Ontological Engineering is a further development of knowledge engineering towards ontology design and creating.

2. Knowledge analysts training outline and organization

The discipline of Knowledge Engineering traditionally emphasized and rapidly developed a range of techniques and tools including knowledge acquisition, conceptual structuring and representation models. These developments have underpinned an emerging methodology that can bridge the gap between the ability of the human brain to structure and store knowledge, and the knowledge engineers' ability to model this process. But for practitioners, knowledge engineering is still a rather new, eclectic domain that draws upon a wide range of areas, including cognitive science, etc. Accordingly, knowledge engineering has been, and still is, in danger of fragmentation, incoherence and superficiality.

Since 2000, a major interest of researchers has focused on building customized tools that aid in the process of knowledge capture and structuring. Trainees are introduced to

major issues in the field and to the role of the knowledge analyst in strategic information system development. We include a lot of interdisciplinary knowledge elicitation and structuring methods that can help the knowledge work, such as the conducting unstructured interview, mastering the verbal reports, business process modelling techniques, road mapping, brainstorming, etc.

The future analysts gain the deep understanding the role of knowledge engineering and knowledge management in companies and organizations; in decision-making by members of an organization; in developing information framework. They study and are trained in practical methods mainly by doing. Attention is given both to developing inter-personal information communication skills and analytical cognitive creative abilities. The first module is targeted at essentials of informal mental modeling by presenting mind maps, concept maps, semantic networks, frames, decision tables, decision trees and other visual forms of knowledge pattern representation.

The training features short lectures, discussions, tests, quizzes and exercises. Lectures are important but the emphasis is put on learning through discussions, simulation, special games, training and case studies. A good deal of the course focuses on auto-reflection and auto-formalizing of knowledge, training of analytical and communicative abilities, discovery, creativity, cognitive styles features, and gaining new insights.

On-the-job or workplace training adds the value of the team spirit and entrained feeling. All the examples are taken just from the every day routine practice. Such approach enables the trainer for better tailoring the course to the specified needs of the company.

Normally the TKE course consists of 4 inter-related modules:

- Getting Started in KE (12 hours),
- Practical KE in depth (12 hours),
- Ontological Engineering (12 hours),
- Business Processes Modeling and mapping (12 hours).

Different combination of sub-topics is possible. Fig.1 illustrates the structure of one variant chosen by Business Engineering Group Company (Saint-Petersburg, Russia).

The main difference of TKE to existing methodologies is cognitive (not technological) bias. The topics of exercises cover categorization, observation, laddering, lateral thinking and other problem solving cognitive methods. Knowledge workers often under-value the significance of psychological background of categorization, laddering and lateral thinking. But during training some of them feel “insight” and become very enthusiastic. We try to implement the ontological approach into the teaching style and strategy. Philosophers of science define ontologism by postulating existence of the systemic hierarchical conceptual specification of any complex object.

Now ontologies help to support knowledge navigation, search and retrieval. They are also used in educational and business research (Blanchard, Mizoguchi, Lajoie, 2009; Dicheva et al, 2005). The practical knowledge workers often underestimate the impact of

their cognitive styles on decision making procedure. Their verbal skills and logics really influence the information processing. It is supposed to be guided by common sense while it needs to be taught and trained.

From organizational point of view the training process consists of series of on-the-job sessions. One group was never not more than 8 persons. Each day classes do not last more then 3-4 hours including the hand-on computer practice in mind-mapping and concept mapping techniques.

3. Teaching ontological thinking and design

Ontologies can be used to describe any business world. But our experience in training shows that nobody can deal with ontologies without knowledge engineering practice. How to teach ontology design? The theory differs from practical need. There are numerous well-known definitions of this milestone term (Gruber, 1993; Guarino and Jiaretta, 1998; Jasper and Uschold, 1999; Mizogushi and Bourdeau, 2000; Neches, 1991) but they may be generalized as "Ontology is a hierarchically structured set of terms for describing an arbitrary domain" (Gomez-Perez et al., 2004). In other words "ontologies are nothing but making knowledge explicit" (Guarino and Welty, 2000).

Since 2000 a major interest of researchers focuses on building customized tools that aid in the process of knowledge capture and structuring. This new generation of tools – such as Protégé, OntoEdit, and OilEd - is concerned with visual knowledge mapping that facilitates knowledge sharing and reuse. The problem of iconic representation has been partially solved by developing knowledge repositories and ontology servers where reusable static domain knowledge is stored. But practitioners from companies and research centres still need simple and constructive algorithms for their activity.

Ontology creating also faces the knowledge acquisition bottleneck problem. The ontology developer encounters the additional problem of not having sufficiently tested practical methodologies, which would recommend what activities to perform. An example of this can be seen when each development team usually follows their own set of principles, design criteria, and steps in the ontology development process. The lack of structured guidelines hinders the development of shared and consensual ontologies within and between the teams. Moreover, it makes the extension of a given ontology by others, its reuse in other ontologies, and final applications difficult (Guarino and Giaretta, 1998; Guarino and Welty, 2000; Jasper and Uschold, 1999).

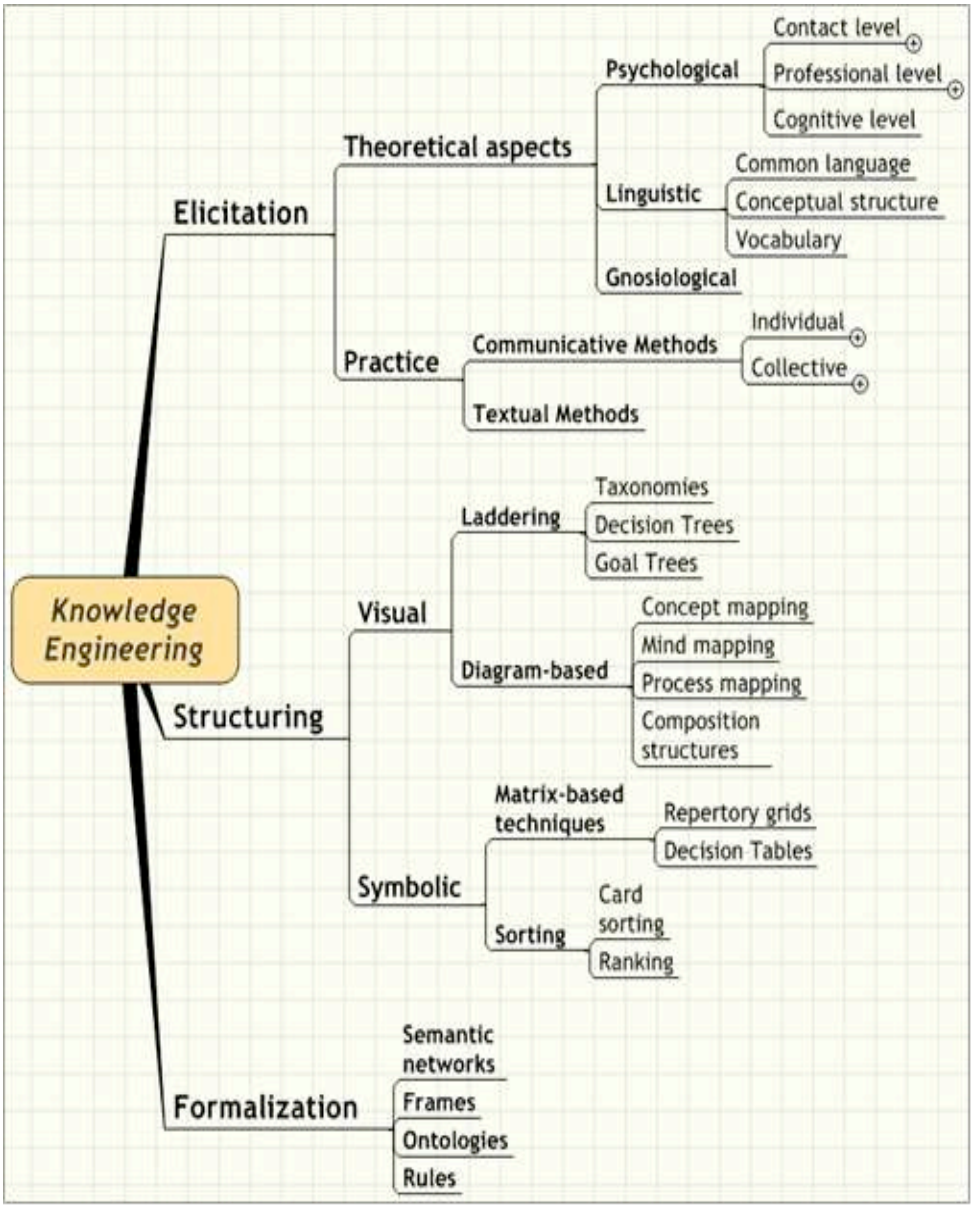

Figure 1. *Outline of training on knowledge engineering*

Several effective methodological approaches have been reported for building ontologies (Swartout, et al 1997; Mizogusgi and Bordeau, 2000; Fridman Noy, Griffin, Musen, 2008).). What they have in common is that they start from the identification of the purpose of the ontology and the needs for the domain knowledge acquisition. However, having acquired a significant amount of knowledge, major researchers propose a formal language expressing the idea as a set of intermediate representations and then generating the ontology using translators. These representations bridge the gap between how people see a domain and the languages in which ontologies are formalized. The conceptual models are implicit in the implementation codes. A re-engineering process is usually required to make the conceptual models explicit.

The idea of using visual structuring of information to improve the quality of user learning and understanding is not new. Concept mapping has been used for more than twenty years (Sowa, 1984; Conlon, 1997; Jonassen, 1998) in system design and development for providing structures and mental models to support the knowledge sharing process. As such, the visual representation of general corporate business concepts facilitates and supports company personnel understanding both substantive and syntactic knowledge. An analyst serves as a knowledge engineer by making the skeleton of the company's data and knowledge visible, and showing the domain's conceptual structure. We try to simplify a bunch of different approaches, terms and notations for practical use and dare to propose a 4-steps recipe for practical ontology design.

3.1. *Ontology design recipe*

The existing methodologies describing ontology life cycle (Uschold and Gruninger, 1996; Mizoguchi and Bourdeau, 2000; Gomez-Perez et al, 2008; Noy, Musen, 2008) deal with general phases and sometimes don't discover the design process in details. Four simple practical steps were proposed in the training course.

Step 1. **Glossary development:** The first step should be devoted to gathering all the information relevant to the described domain. The main goal of this step is selecting and verbalizing all the essential objects and concepts in the domain.

Step 2. **Laddering:** Having all the essential objects and concepts of the domain in hand, the next step is to define the main levels of abstraction. It is also important to elucidate the type of ontology according to ontology classification, such as taxonomy, partonomy, or genealogy.

This is being done at this step since it affects the next stages of the design. Consequently, the high level hierarchies among the concepts should be revealed and the hierarchy should be represented visually on the defined levels.

Step3. **Disintegration and Categorization:** the main goal of this step is breaking high level concepts, built in the previous step, into a set of detailed ones where it is needed. This could be done via a top-down strategy trying to break the high level concept from the root of previously built hierarchy. At the same stage, detailed concepts are revealed in a structured hierarchy and the main goal at this stage is generalization via bottom-up structuring strategy. This could be done by associating similar concepts to create meta-concepts from leaves of the aforementioned hierarchy.

Step 4. **Refinement**: The final step is devoted to updating the visual structure by excluding the excessiveness, synonymy, and contradictions. As mentioned before, the main goal of the final step is try to create a beautiful ontology. We believe what makes ontology beautiful is harmony.

Using these tips the trainees developed several huge company ontologies (Gavrilova, Laird, 2005).

3.2. *"Beatification" of business ontology*

The idea of the good shape in modelling is rather common in science. Let's try to apply this approach to the ontology design. One of substantial impulse to it was given by German psychological school of M. Wertheimer. His idea of good Gestalt (image or pattern) may be transferred into ontological engineering design guidelines. Some essential Gestalt principles of this school (Werthheimer, 1959):

- Law of Pragnanz (M. Wertheimer) - organization of any structure in nature or cognition will be as good (regular, complete, balanced, or symmetrical) as the prevailing conditions allow (law of good shape).
- Law of Proximity – objects or stimuli that are viewed being close together will tend to be perceived as a unit.
- Law of Similarity – things that appear to have the same attributes are usually perceived as being a whole.
- Law of Inclusiveness (W.Kohler) - there is a tending to perceive only the larger figure and not the smaller when it is embedded in a larger.
- Law of Parsimony – the simplest example is the best or known as Ockham's razor principle (14-th century): ``entities should not be multiplied unnecessarily".

We suggest to use these laws for pursuing conceptual balance and clarity of corporate knowledge ontology.

3.2.1 Conceptual balance

A well-balanced ontological hierarchy equals a strong and comprehensible representation of the domain knowledge. However, it is a challenge to formulate the idea of a well-balanced tree. Here we offer some tips to help formulate the "harmony":

- Concepts of one level should be linked with the parent concept by one type of relationship such as is-a, or has part.
- The depth of the branches should be more or less equal (±2 nodes).
- The general outlay should be symmetrical.
- Cross-links should be avoided as much as possible.

3.2.2 Clarity

Moreover, when building a comprehensible ontology it is important to pay attention to clarity. Clarity may be provided through number of concepts and type of the relationships among the concepts. Minimizing the number of concepts is the best tip according to Law of Parsimony. The maximal number of branches and the number of levels should follow

Miller's magical number (7±2) (Miller, 1956). Furthermore, the type of relationship should be clear and obvious if the name of the link is missed.

At the first stages it is possible to use any of the available graphical editors to design an ontology, e.g. PaintBrush, Visio, Inspiration. A nice layout can be reached by using mindmapping tools as Freemind™, MindManager™ or Visual Mind ™. The trainees really enjoyed the process of "beatification" of their ontologies during test exercises.

As an example we may discuss the ontology presented at Fig.2. This figure maps the ontology of knowledge engineering. We try to follow all the rules described earlier, but one can see that the branch "formalization" is too short and shallow. In our case it is understandable because of the specifics of the course which was aimed at non-programmers. But in the general ontology of this field should be detailed better.

4. Discussion

Challenges have fueled opportunities for analytic tool developers, educators, and business process owners that support analytic communities in the management of knowledge, information and data sources. The field of Knowledge Management has undergone several bouts of high hopes and press-induced hype ending with grave disappointment and missed promises. All too often we see old Information Management technology repackaged and retagged as the latest KM offering. However, today we still have functioning corporate KM systems being arranged by qualified knowledge workers. It is expected that large corporations will be forced to rethink their knowledge management strategies towards the human factors assessment. We hope that training and couching of knowledge analysts will rise the new types of business development platforms and will play a key role in the articulation of the corporate KM landscape in the next 3-5-7 years.

Any mature company needs business analysts. Analysts are super-knowledge workers, but even they enter "the world of ontologies" with some doubt. But in the training their interest grows and rather soon they begin to use ontologies in their practical work. Our experience in training of knowledge analysts in the period of 1999-2010 confirm the unique role of knowledge structuring for developing ontologies quickly, efficiently and effectively. We follow David Jonassen's idea of using concept maps as "a mind tool" (Jonassen,1998). The use of visual paradigm for the representing and supporting the training process not only helps a professional trainer to concentrate on the problem rather than on details, but also enables students to process and understand greater volume of information. After training major of the trainees were able to map their professional knowledge using different visual forms of ontology design – from mind maps to concept maps. They developped the ontologies of the customers, suppliers, products, solutions, requirements, projects, etc.

Business is based on knowledge processing in new information age. So the skillfull knowledge workers can really increase the productivity and sustainability of modern

business practice in the innovate service-oriented economy. And the use and development of ontologies help to annotate information so that diverse groups of humans and machines can process it more meaningfully.

Acknowledgements

The work was partially funded by grantы of Russian Foundation of Basic Research and grant from St.Petersburg State University. Thanks to all of my students and trainees whose questions and discussion help a lot to the improvement of the course.

References

Blanchard E., Mizoguchi R., Lajoie S. (2009) *Addressing the Interplay of Culture and Affect in HCI: An Ontological Approach* //Human-Computer Interaction. Ambient, Ubiquitous and Intelligent Interaction, 13th International Conference, HCI International 2009, USA, July 19-24, 2009, Proceedings, Part III. Lecture Notes in Computer Science 5612 Springer, 575-584

Boose, J.H. (1990) *Knowledge Acquisition Tools, Methods and Mediating Representations. In Knowledge Acquisition for Knowledge-Based Systems* (Motoda, H. et al., Eds), IOS Press, Ohinsha Ltd., Tokyo, pp.123-168.

Dicheva D., Gavrilova, T., Sosnovsky S., Brusilovsky P. (2005) Ontological Web Portal for Educational Ontologies. In Proc. Of "*Applications of Semantic Web Technologies for E-Learning Workshop (SW-EL'05)"* in conjunction with 12th Int.Conf. on Artificial Intelligence in Education (AI-ED'05), Amsterdam, 19-29.

Conlon T. (2000) *'Visions of Change: Information Technology, Education and Postmodernism* // British Journal of Educational Technology Vol 31 No 2, pp. 109-116.

Firestone J., McElroy M. (2005) *Doing Knowledge Management* // The Learning Organization, 12, no. 2 (April, 2005), pp.189-212 .

Firestone J. (2003) *Enterprise Information Portals and Knowledge Management*, Burlington, MA: KMCI Press/Butterworth-Heinemann.

Fridman Noy N., Griffith N., Musen M. (2008) *Collecting Community-Based Mappings in an Ontology Repository*, In International Semantic Web Conference, pp. 371-386.

Jasper, R. and Uschold, M (1999*). A Framework for Understanding and Classifying Ontology Applications*. In 12th Workshop on Knowledge Acquisition Modelling and Management KAW'99.OilEd (2004) Bechhofer, S. and Ng G. http://oiled.man.ac.uk/

Jonassen, D.H. (1998) *Designing constructivist learning environments. In Instructional design models and strategies* (Reigeluth, C.M. (Ed), 2nd ed., Lawrence Eribaum, Mahwah, NJ.

Gavrilova, T., Voinov, A. (1998) *Work in Progress: Visual Specification of Knowledge Bases* // Lecture Notes in Artificial Intelligence 1416 "Tasks and Methods in Applied Artificial Intelligence", A.P.del Pobil, J.Mira, M.Ali (Eds), Springer, pp. 717-726.

Gavrilova, T.A., Voinov, A., Vasilyeva E. (1999) *Visual Knowledge Engineering as a Cognitive Tool* // Lecture Notes in Artificial Intelligence, Proc. of Int. Conf. on Artificial and Natural Networks IWANN'99, Spain, Benicassim, pp.123-128.

Gavrilova, T. (2003) *Teaching via Using Ontological Engineering* // Proceedings of XI Int. Conf. "Powerful ICT for Teaching and Learning" PEG-2003, St.Petersburg, pp. 23-26.

Gavrilova, T., Kurochkin M., Veremiev V. (2004) *Teaching Strategies and Ontologies for E-learning* // Int. J. "Information Theories and Applications", vol.11, N1, pp.35-42.

Gavrilova, T., Laird, D. *Practical Design of Business Enterprise Ontologies* (2005) In Industrial Applications of Semantic Web (Eds.Bramer M., Terzyan V., Springer, pp.61-81.

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

Gruber, T. (1993) *A translation approach to portable ontology specifications*. Knowledge Acquisition , Vol. 5, pp.199- 220.

Guarino, N. & Giaretta, P. (1998) *Ontologies and Knowledge Bases: Towards a Terminological Clarification.* // Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing, IOS Press, pp.25- 32.

Guarino, N., Welty, C. (2000) *A Formal Ontology of Properties*. In R. Dieng and O. Corby (eds.), Knowledge Engineering and Knowledge Management: Methods, Models and Tools. 12th International Conference, EKAW2000. Springer Verlag, pp. 97-112.

Leydesdorff L. (2006) *The Knowledge-Based Economy Modeled, Measured, Simulated.* Universal-Publishers/.

Miller, G. (1956) *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information* // The Psychological Review, v. 63, pp. 81-97.

Mizogushi, R. and Bourdeau J. (2000) *Using Ontological Engineering to Overcome Common AI-ED Problems*. International Journal of Artificial Intelligence in Education, v.11, pp.1-12.

Neches, et al (1991) *Enabling Technology for Knowledge Sharing*. AI Magazin, Winter, pp.36- 56.

Protégé (2007), Stanford Medical Informatics. Accessed from http://protege.stanford.edu/

Scott, A., Clayton, J.E. & Gibson E.L. (1994) *A Practical Guide to Knowledge Acquisition*, Addison-Wesley.Swartout, B., Patil, R., Knight, K. & Russ, T. (1997) Toward Distributed Use of Large-Scale Ontologies. In Ontological Engineering, AAAI- 97 Spring Symposium Series, pp.138- 148.

Sowa, J. F. (1984) *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley, Reading, Massachusetts.

The CIO's Guide to Semantics (2004) Semantic Arts©, Inc. www.semantic-conference.com

Tu, S., Eriksson, H., Gennari, J., Shahar, Y. & Musen M. (1995) *Ontology-Based Configuration of Problem-Solving Methods and Generation of Knowledge-Acquisition Tools*. In "Artificial Intelligence in Medicine", N7, pp.257-289.

Werthheimer, M. (1959) *Productive Thinking*, HarperCollins/

Wielinga, B., Schreiber, G. & Breuker J. (1992) *A Modelling Approach to Knowledge Engineering*. In Knowledge Acquisition, 4 (1), Special Issue, pp.23-39.

Wiig E.H & Wiig K. M. (1999) *On Conceptual Learning Knowledge Research Institute*, Inc. Working Paper 1999-1. http://www.krii.com/downloads/conceptual_learning.pdf

Uschold, M., Gruninger M (1996*). "Ontologies: Principles Methods and Applications"*, Knowledge Engineering Review, v.1, N1.