

PRACTICAL DESIGN OF BUSINESS ENTERPRISE ONTOLOGIES

Tatiana Gavrilova; David Laird

*Intelligent Computer Technologies Dept. Saint-Petersburg State Polytechnic University¹;
School of Information Science, University of Pittsburgh²*

Abstract: This paper presents one approach for developing enterprise ontologies. The underlying research framework is pursuing a methodology that will aid the process of knowledge structuring and practical ontology design, with emphasis on visual techniques. For illustration of the proposed technique, the development of a practical ontology of information technology skills for a human resources knowledge management system is described.

Key words: Ontology, Visual Knowledge Engineering, Knowledge Acquisition, Knowledge Management

1. INTRODUCTION

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. 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. 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.

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^{1,2,3} 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 of 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.

At the present time, this structure is called an ontology. However, ontology-based approaches to business are relatively new and fertile research areas. They originated in the area of knowledge engineering^{4,5}, then evolved into ontology engineering^{6,7}.

The discipline of Knowledge Engineering traditionally emphasized and rapidly developed a range of techniques and tools including knowledge acquisition, conceptual structuring and representation models^{8,9}. 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. This new generation of tools – such as Protégé, OntoEdit, and OilEd - is concerned with visual knowledge mapping that facilitates knowledge sharing and reuse^{10,11,12}. The problem of iconic representation has been partially solved by developing knowledge repositories and ontology servers where reusable static domain knowledge is stored. Ontolingua, Ontobroker and many others are examples of such projects^{13,14}.

The usage of ontologies has special value in companies where specialists reuse domain ontologies in order to support the business protocols that are grounded in the domain's problem-solving methodology. Therefore, the basic idea is to allow experts to model both domain and problem-solving knowledge using the same visual language. Knowledge entities that represent static knowledge of the domain are stored in hierarchical order in the knowledge repository and can be reused by others. At the same time, those knowledge entities can also be reused in description of the properties or methodological approach as applied in the context of another related knowledge entity. The involved concept map modeling language is based upon a class-based object-oriented language that supports the classification and parameterization of knowledge entities.

This paper proposes a practical approach to business ontology design. The underlying research is pursuing usage of visual iconic representation

and diagrammatic structures, with emphasis on visual design. For clearer understanding of the methodology, the process of developing a practical ontology of information technology knowledge and skills is described. In the remainder of the paper, we will describe some theoretical issues regarding ontological engineering and present our proposed methodology for ontology design. Moreover, we will describe our detailed practical example using the proposed methodology. In conclusion, we provide insight through discussion of current and possible future work.

2. USING ONTOLOGICAL ENGINEERING FOR BUSINESS APPLICATIONS

We start the discussion of theoretical issues of ontological engineering by developing a definition of ontology from literature within the field.

2.1 Ontology Definition

Ontology is a set of distinctions we make in understanding and viewing the world. There are numerous well-known definitions of this seminal term^{15,16,17}, that may be generalized by such:

“Ontology is a hierarchically structured vocabulary describing a domain that can be used as a skeletal foundation for a knowledge base”.

This definition clarifies the ontological approach to knowledge structuring while providing sufficient freedom for open-ended, creative thinking. For example, ontological engineering can provide a clear representation of a company’s structure, human resources, physical assets, and products, and their inter-relationships. Many researchers and practitioners argue about the distinctions between the ontology and the user’s conceptual model. We believe that the ontology corresponds to the analyst’s view of the conceptual model, but is not the *de facto* model.

Ontology as a useful structuring tool may greatly enrich the business modeling process, providing users of KM-systems an organizing axis to help them mentally mark their vision of the domain knowledge.

2.2 Creating Ontologies for Business Use

Ontology creation faces the knowledge acquisition bottleneck problem. The ontology developer frequently encounters the additional problem of not utilizing sufficiently tested and generalized methodologies, which would recommend what activities to perform and at what stage of the ontology development process. An example of this can be seen when each

development team generally follows its own set of principles, design criteria, and steps in the ontology development process. The lack of structured guidelines and methods hinders the development of shared and consensual ontologies within and between the teams. Moreover, it makes the extension of a given ontology by others and its reuse in other ontologies and final applications difficult¹⁸.

Several effective domain-independent methodological approaches have been reported for building ontologies^{6,7,19}. What these approaches have in common is that they consistently begin with identification of the purpose of the ontology, and the need for acquisition of the domain's knowledge. 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. Ontological commitments and design criteria are implicit in the ontology code.

Figure 1 presents our vision of the mainstream state-of-the-art categorization in ontological engineering^{20,21,22} and may help the knowledge analyst to figure out what type of ontology he/she really needs. We use Mindmanager™ as it proved to be a powerful visual tool.

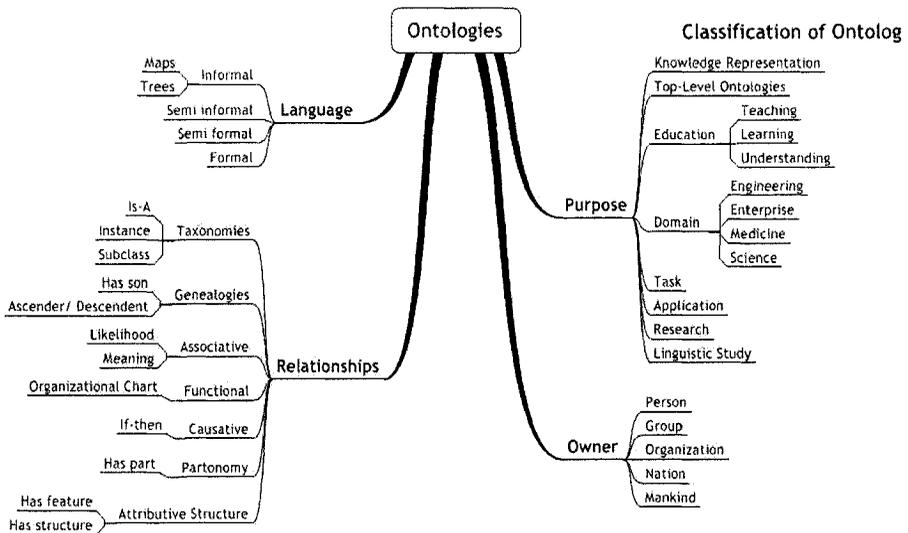


Figure 1. Ontology classification

Frequently, it is impossible to express company business information in a single ontology. Accordingly, company knowledge storage consists of a set of related ontologies. However, some problems may occur when moving from one ontological space to another that could be solved by constructing meta-ontologies that may help to resolve these problems.

We can propose different types of ontologies that can support business applications:

- Company organizational structure
- Main concepts ontology (products, services, customers, skills, etc.),
- Historical ontology (genealogy of owners, customers, products, services, etc.),
- Partonomy of the company knowledge
- Taxonomy (methods, techniques, technologies, business-processes, skills, etc.)

The concrete set of ontologies depends on personal vision, business application and awareness level of the system's analysts and users. Generalizing our experience in developing different business and teaching ontologies in the field of consulting, business modeling and information technologies^{23,24,25,26}, we propose a four-step algorithm that may be helpful for visual ontology design. We try to develop the ideas of Uschold and King's skeletal methodology²² putting stress on details of ontology capture, where visual representation works as a powerful mind tool² in the structuring process. Visual form influences both analyzing and synthesizing procedures in ontology development process. That is why we believe that the "beauty" of the ontology plays an important role in understanding of the knowledge.

3. ONTOLOGY CREATING

While in major works the emphasis is put on ontology specification, we would like to elucidate the essentials of ontology capture²², not coding.

3.1 Four-Step Algorithm

Step1. Goals, strategy and boundary identification: The first step in ontology development should be to identify the purpose of the ontology and the needs for the domain knowledge acquisition. It is important to be clear about why the ontology is being built and what its intended uses are²². We also need to define the scope or "boundaries" of the ontology, before

compiling a glossary. It is also important to elucidate the type of ontology according to Figure 1 classification, such as taxonomy, partonomy, and genealogy. That effort is done at this step, as it affects the next stages of the design.

Step2. Glossary development or meta-concept identification: This time consuming step is devoted to gathering all the information relevant to the described domain. The main goal of this step is selecting and verbalizing all of the essential objects and concepts in the domain. A battery of knowledge elicitation techniques may be used – from interviews to free association word lists.

Step3. Laddering, including categorization and specification: Having all the essential objects and concepts of the domain in hand, the next step is to define the main levels of abstraction. Consequently, the high level hierarchies among the concepts should be revealed and the hierarchy should be represented visually on the defined levels. This could be done via a top-down strategy by trying to break the high level concept from the root of the previously built hierarchy, by detailing and specification of instance concepts. Revealing a structured hierarchy is one of the main goals at this stage. Another way is generalization via bottom-up structuring strategy. Associating similar concepts to create meta-concepts from leaves of the aforementioned hierarchy could do this. The main difficulty is forming categories by creating high level concepts and/or breaking them into a set of detailed ones where it is needed.

Step4. Refinement: The final step is devoted to updating the visual structure by excluding any excessiveness, synonymy, and contradictions. As mentioned before, the main goal of the final step is try to create a beautiful ontology. The ideas of “beatification” are well known in basic studies beginning from the search for beautiful formula, model or result. Beauty was always a very strong criterion of scientific truth. We believe that harmony and clarity are what make an ontology beautiful.

3.2 Harmony

To achieve harmony, we attempt to follow Gestalt (good form) principles by M. Wertheimer²⁷:

- Law of Pragnanz: 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 tendency to perceive only the larger figure, and not the smaller, when it is embedded in a larger.
- Law of Parsimony – the simplest example is best known as Ockham's razor principle (14th century): "entities should not be multiplied unnecessarily".

3.2.1 Conceptual balance

A well-balanced ontological hierarchy equates to a strong and comprehensible representation of the domain's knowledge. Ill-balanced ontology design (at Figure 2) shows that long branches are over-detailed, while shorter ones are under-investigated. For our problem, this may create a situation where some IT skills will be described too precisely, while others will be just briefly mentioned. Ill-balanced ontology often demonstrates the low professional level of the expert and/or knowledge analyst. 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 at one level should be linked with the parent concept by only 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.

Fig.2 illustrates the balance idea.

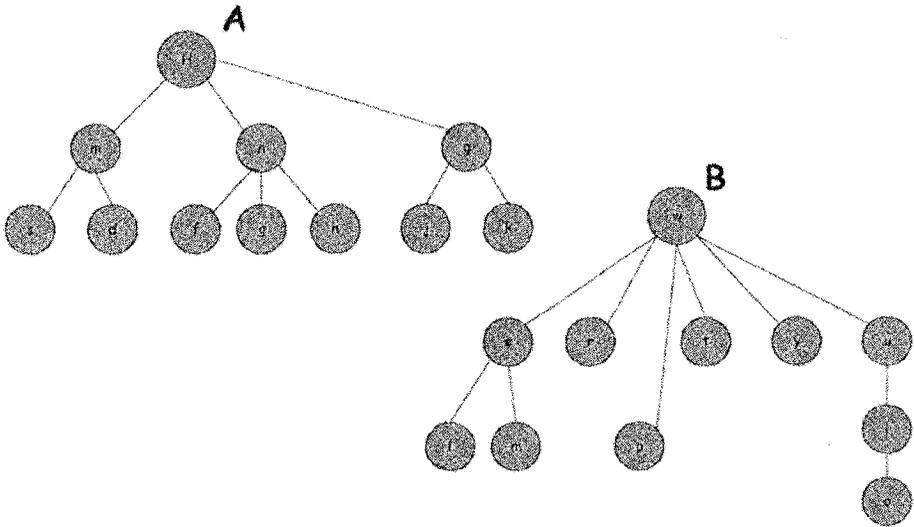


Figure 2. Well-balanced (A) and ill-balanced (B) ontologies

3.3 Clarity

In addition to the principle of harmony, it is important to pay attention to clarity when building a comprehensible ontology. Clarity may be provided through a number of concepts, and types of the relationships among the concepts.

- Minimizing the number of concepts. The maximal number of branches and the number of levels should follow Miller's magical number $(7 \pm 2)^{28}$.
- Furthermore, the type of relationship should be clear and obvious if the name of the relationship is missing.

Some tips to achieve visual clarity are described later in section 4.4.

4. DEVELOPING A PRACTICAL ONTOLOGY

In this section we describe the development of an ontology of information technology skills and knowledge, following the aforementioned 4-step algorithm. We have tried to report the exact practical procedures we followed at each step by including all the visual structures.

4.1 Step 1 - Purpose and use of Ontology:

It is important to first identify the purpose and proposed usage of the ontology early in the process of its development²². The example ontology described throughout the remainder of this paper was developed to support a business application to address the following needs.

Situation/Problem: A company is seeking to identify the knowledge and skills of each of its employees that are relevant to the work of the company. This data will allow the company to:

- Identify the essential skills of the organization
- Develop a knowledge retention strategy to ensure that sufficient depth is present in the organization in the event of resignation, retirement, or other loss of key employees
- More effectively identify and utilize employee skillsets:
 - Allow employees to quickly find experts to address unique questions or problems.
 - Identify individuals in the company with the needed skill to work on new or expanding projects
- Develop individual and organization-wide training plans and strategy, based on the collective training needs of the enterprise.

Solution: Use a network-based intranet application that allows employees to identify their individual skills and training needs. This application will make use of an ontology of skills that span the IT industry, and allow employees to select relevant skills and knowledge from that ontological presentation of skills, which they currently possess, or have a business need to acquire. Use of the ontology in this way serves the following purposes:

- Ensures that each employee considers the entire range of IT Skills that he might possess, or that are relevant to the organization.
- Ensures that data is entered uniformly into the system by each employee, with a consistent understanding of the meaning of each skill. This consistency allows subsequent searches of the employee skills database to find all cases of a selected skill, and the organization's training to be planned for specific or broad categories of skills.
- Provides for a framework to visualize and better understand the relationships of skills that are relevant and critical to the success of the organization.

4.2 Step 2 - Glossary Development

As previously mentioned, the first step in building an ontology is collecting information in the domain and building a glossary of the terms of the domain. To build a glossary of information technology skills, we collected the terms from two different types of resources: closed-corpus material and open-corpus material.

The closed corpus materials are in the form the company's job descriptions in the field of information technology, recent correspondence and status reports from the Information Systems department, and an organizationally generated skills inventory. The open corpus materials include the table of contents and index of general information systems text references, categorizations and descriptions of computer and information science course offerings, and existing ontologies about the field of information technology. The terms and concepts from each of these sources were combined to build a single glossary.

Table 1. Sampling of a Glossary of IT Skills and Knowledge

Personal Computer Maintenance	Peripherals Maintenance	Help Desk Support
Database Administration	Programming	Application Development
Cyber Security	Encryption	Commercial Software
Super Computing	Telecommunications	Training
Project Management	Graphics	Mobile Computing
Computer Architecture	Software Development Lifecycle	Quality Assurance
Human Factors in Systems	Human Computer Interactions	Artificial Intelligence
Geographic Information Systems	Decision Support Systems	Data Mining
Information Storage and Retrieval	Programming Languages	Software Engineering
Algorithm Design	Computer Engineering	Visual Languages
Operating Systems	Document Processing	Information Processing Standards
Knowledge Representation	Legal and Ethical Issues	Expert Systems
Ontologies	Knowledge Management	Routers
Bridges	Network Switches	Computer Server Support
Virus Detection	Email Systems	Enterprise System Customization

4.3 Step 3 - Laddering: Building an Initial Mind Map Structure

In the third step, we built an initial visual structure of the glossary terms. The main goal of this step is the creation of a set of preliminary high level concepts and the categorization of the glossary terms into those concepts. A mind map can be a useful visual structure for this step. **Figure 3** presents the mind map of our initial categorization. Since the categorization in this step is preliminary, some of terms might not fit into any of the initial categorization. We should mention that the categorization in this step is done entirely manually. However, we employed the job descriptions, text glossary and table of contents, and groupings of university course offerings, which were used to build the glossary in the previous step, to build the initial categories as well. We can consider the groupings from these sources to be expert help in designing the ontology, because such groupings were accomplished to make the presented information in these sources clear and easily accessible; traits that we desire in the finished IT Skills ontology.

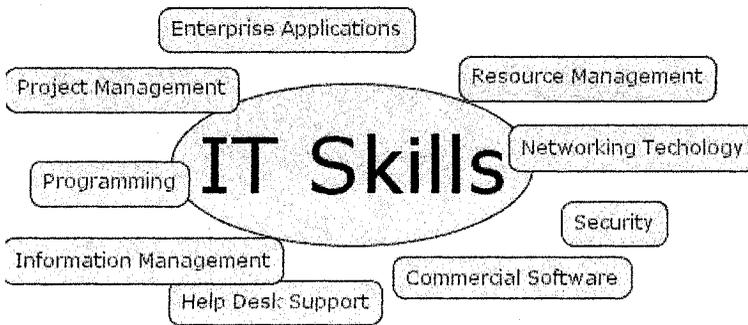


Figure 3. Trivial Categorization

Figure 4 presents the details of our initial categorization of the terms into the concept in **Figure 3**. The visual structures presented in this step illustrate the idea of how an ontology can bridge the gap between the chaos of unstructured data presented in the glossary, and be a clear means of showing mapped representations.

Later, we composed more precise concepts and hierarchies by analyzing the glossary and previously built visual structure. First we employed the top-down design strategy to create meta-concepts such as Programming, Network Support, Project Management, etc. Then using the bottom-up strategy we tried to fit the terms and concepts into the meta-concept. Moreover, we created the relationships between the concepts. A concept map is the most useful visual structure for representation of the results of this stage, since it gives the ability of defining the relationship in addition to building the hierarchy. The output of this step is a large and detailed map, which covers the domain hierarchically.

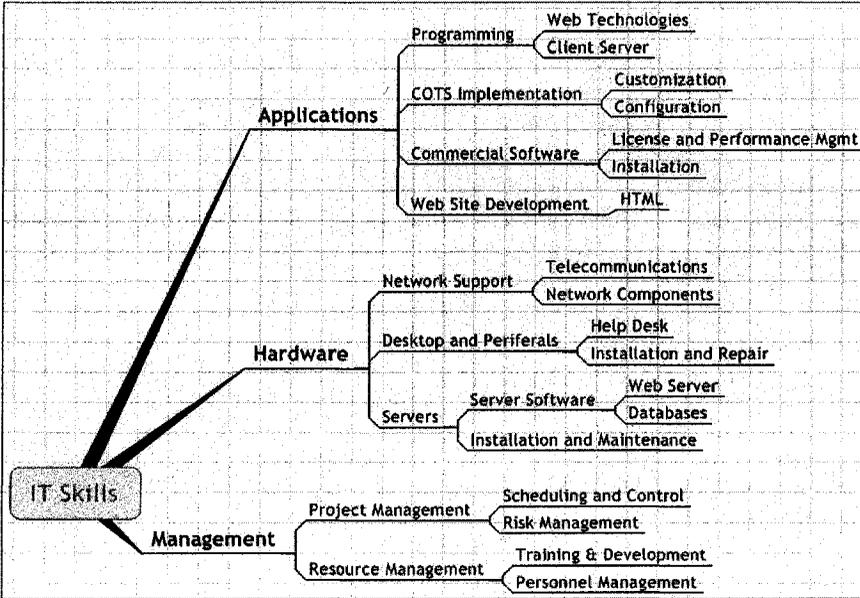


Figure 4. Details of first level categorization

Next, based on the detailed concept map, we built the general ontology that is shown in **Figure 5**, utilizing liberal relationship terms to link the concepts with detailed terms from the glossary.

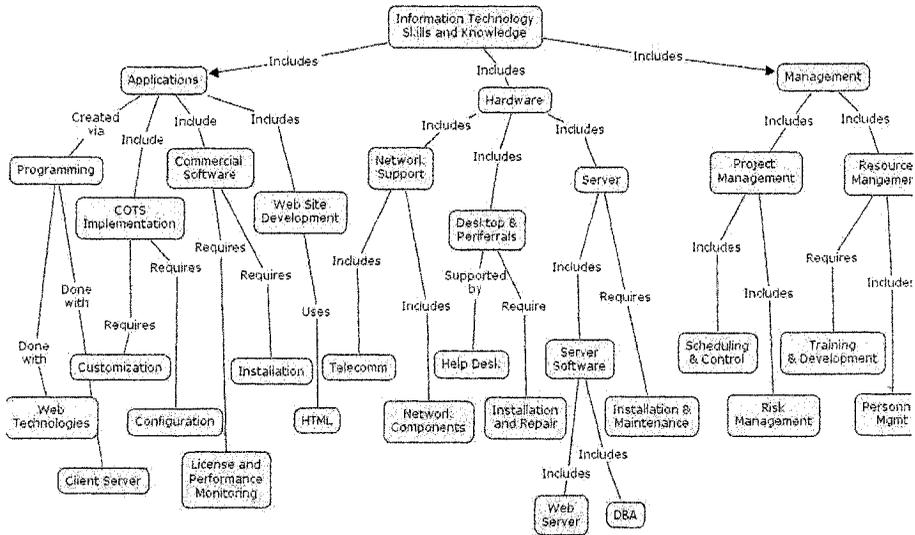


Figure 5. General ontology

4.4 Step 4: Refinement

As described in the algorithm, the final step is devoted to making the ontology beautiful. The following are some practical tips that may be taken into consideration while refining the ontology, and are illustrated in Figure 6:

1. Use different font sizes for different strata
2. Use different colors to distinguish particular subsets or branches (not very clear in the black and white printout).
3. Use a vertical layout of the tree structure/diagram.
4. If needed, use different shapes for different types of nodes.

Moreover, we re-built the general ontology while taking into consideration the harmony and clarity factors. Comparing **Figure 5** and **Figure 6** presents these changes. Another feature of harmony is having the same relationship in every level. Moreover, to achieve clarity, we removed all unnecessary nodes and use standard, consistent relationships to simplify understanding.

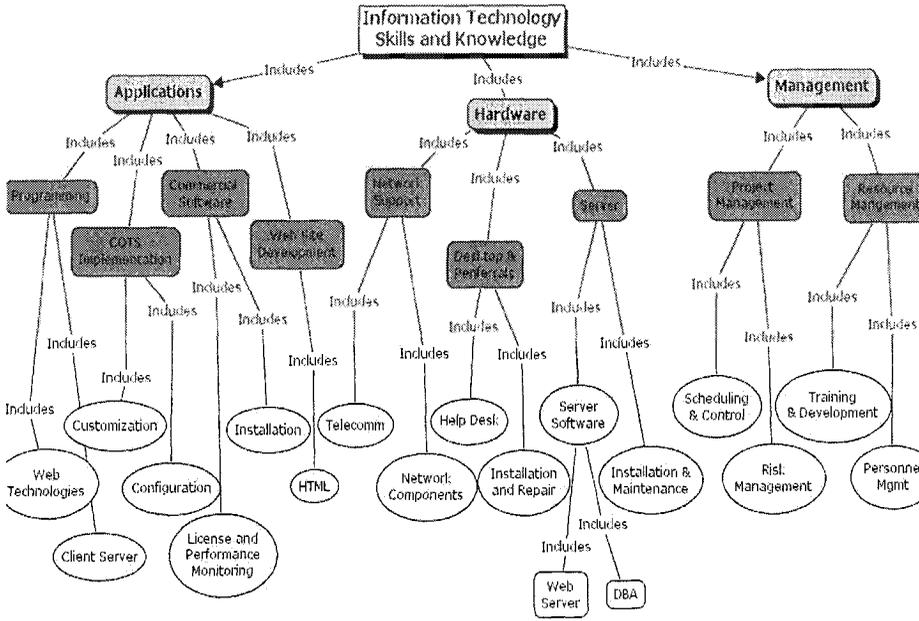


Figure 6. Harmony and clarity in the ontology

5. DISCUSSION

Our research stresses the role of knowledge structuring for developing ontologies quickly, efficiently, and effectively. At a basic level of knowledge representation, within the context of everyday heuristics, it is easier for practitioners simply to draw the ontology using conventional “pen and pencil” techniques. However, for more sophisticated knowledge representations, our proposed 4-step ontology development process is proposed.

Development and use of an ontology of IT Skills and Knowledge was illustrated in this paper to provide a concrete example of the proposed methodology. A more detailed version of the illustrated ontology will be integrated into the use of a Knowledge Management application, used to develop a map of critical skills and knowledge within a business enterprise. This awareness of the critical skills needed and possessed by the

organization will allow strategies to be developed to ensure the retention and most effective use of those critical skills. Without a comprehensive ontology to frame this investigation, valid and useful results would be far more difficult to achieve.

In subsequent research, we plan to explore ways that ontology development and use can further improve visualization of business needs, and deliver additional value to the organization. Such investigations will address the reduction of overlap between business units in an organization, aligning recruiting efforts with actual business needs, development of job descriptions that accurately reflect the skills and knowledge truly needed for the success of the organization, and clearer understanding of the most critical and valued skills within the organization.

6. ACKNOWLEDGMENTS

This work is partly supported by the grant of Russian Foundation for Basic Research No.04-01-00466.

7. REFERENCES

1. J.F. Sowa (1984) *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley, Reading, Massachusetts.
2. D.H. Jonassen (1998) *Designing constructivist learning environments*. In *Instructional design models and strategies* (C.M. Reigeluth (Ed), 2nd ed., Lawrence Erlbaum, Mahwah, NJ.
3. T. Conlon (1997) *Towards Diversity: Advancing Knowledge-based Modelling with Knowledge Acquisition*. In *Proceedings of 8th International PEG Conference*, Sozopol, Bulgaria., pp.379-386.
4. J.H. Boose (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.
5. P. Brusilovsky and R. Rizzo (2002), *Map-Based Horizontal Navigation in Educational Hypertext*. In *Proceedings of Hypertext*, University Of Maryland, College Park, USA
6. D. Fensel (2001) *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce*. Springer.
7. R. Mizogushi and J. Bourdeau (2000), *Using Ontological Engineering to Overcome Common AI-ED Problems*. *International Journal of Artificial Intelligence in Education*, volume 11, pp.1--12.
8. H. Adeli (1994) *Knowledge Engineering*. McGraw-Hill, New York.
9. A. Scott, J.E. Clayton & E.L. Gibson (1994) *A Practical Guide to Knowledge Acquisition*, Addison-Wesley.
10. Protégé, Stanford Medical Informatics. Accessed from <http://protege.stanford.edu/> at December 07, 2004
11. OntoEdit, AIFB, University of Karlsruhe. Accessed from

- <http://www.ontoknowledge.org/tools/ontoedit.shtml> at December 07, 2004
12. OilEd, S. Bechhofer and G. Ng Accessed from <http://oiled.man.ac.uk/> at December 7, 2004
 13. OntoBroker, Accessed from http://ontobroker.aifb.uni-karlsruhe.de/index_ob.html at December 7, 2004
 14. Ontolingua, Stanford University. Accessed from <http://www.ksl.stanford.edu/software/ontolingua/> at December 7, 2004
 15. Neches, et al , (1991) Enabling Technology for Knowledge Sharing. AI Magazin., Winter, pp.36- 56.
 16. T. Gruber (1993) A translation approach to portable ontology specifications. Knowledge Acquisition , Vol. 5, pp.199- 220.
 17. A. Gómez-Pérez, M. Fernández-López, O. Corcho (2004) Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web Springer.
 18. N. Guarino & P. Giaretta (1998) Ontologies and Knowledge Bases: Towards a Terminological Clarification. In Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing, IOS Press, pp.25- 32.
 19. B. Swartout, R. Patil, K. Knight & T. Rus (1997) Toward Distributed Use of Large-Scale Ontologies. In Ontological Engineering, AAAI- 97 Spring Symposium Series, pp.138-148.
 20. N. Guarino, C. Welty (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: 97-112.
 21. R. Jasper and M. Uschold (1999). A Framework for Understanding and Classifying Ontology Applications. In Twelfth Workshop on Knowledge Acquisition Modeling and Management KAW'99.
 22. M. Uschold, M. Gruninger (1996). Ontologies: Principles Methods and Applications. In Knowledge Engineering Review, Volume 11, Number 2.
 23. T.A. Gavrilova & A. Voinov (1996) Visualized Conceptual Structuring for Heterogeneous Knowledge Acquisition . In Proceedings of International Conference on Educational Multimedia and Hypermedia, EDMEDIA'96, MIT, Boston, USA, pp.258-264.
 24. T. Gavrilova, A. Voinov (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.
 25. T.A. Gavrilova, A. Voinov, E. Vasilyeva (1999) Visual Knowledge Engineering as a Cognitive Tool, Proc. of Int. Conf. on Artificial and Natural Networks IWANN'99, Spain, Benicassim. - pp.123-128.
 26. T. Gavrilova (2003) Teaching via Using Ontological Engineering, Proceedings of XI Int. Conf. "Powerful ICT for Teaching and Learning" PEG-2003, St.Petersburg. – p. 23-26.
 27. M. Wertheimer (1982) Productive Thinking. Univ of Chicago Pr (Tx); Phoenix ed edition,
 28. G. Miller (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review, vol. 63, pp. 81-97.