Open-source system based on the Ontopia engine

PETR DO, LADISLAV BURITA
CIS Department, Faculty of Military Technology
University of Defence
Kounicova 65, Brno, 662 10
CZECH REPUBLIC
petr.do@unob.cz, ladislav.burita@unob.cz

Abstract: The article is based on the doctoral thesis – Knowledge management in military area and the defence research project – Knowledge management of the NEC in the Army of the Czech Republic – MENTAL. These works cover theoretical part of knowledge management area and the research which is focused on modelling and development concept of knowledge-based systems. The main technology base is standard ISO – Topic Maps (TM) which is used in application environment Ontopia. This tool represents the logical layer of knowledge-based systems and also which interprets knowledges.

Key-Words: AToM2, knowledge management, knowledge systems, MENTAL, ontology, Ontopia, WebAppOnt.

1 Introduction
The article describes the experience on the doctoral thesis and the defence research project. It is focused on the procedure from the selection of terms from specific domain to design domain ontology. The paper next explains the ways reusing of an existing taxonomy in an ontology construction, presents the possibilities of using the taxonomy built in the selected domain when creating its domain ontology, explains the difference between taxonomies and ontologies in various contexts, and focuses on the description of the specific domain ontology and the use of the existing taxonomy for its building.

Both works are situated to Knowledge Management (KM) area which could be implemented in the Army of the Czech Republic (ACR). The aim of the article is to carry out the analysis of knowledge approaches, ontologies and ontology languages, and to assess their suitability for using them in the ACR; to propose a methodology for knowledge systems development in the ACR; to elaborate a knowledge system proposal in the ACR and to implement it. The ideological models of our research have become portals of the U.S. Army (AKO, DKO).\footnote{AKO, DKO, 2010, www.us.army.mil}

The accomplishment of the project is assured by successful cooperation of researchers from the University of Defence (UoD) and the huge developers support for the Ontopia tool. The most important activity concerning the knowledge-based system (KBS) is the design and development of an appropriate ontology, which constitutes a formal framework for storing the knowledge, creating links between knowledge and ontology concepts, and establishing connections to concepts and pieces of knowledge of vital documents, which are connected with the area in focus. Ontology itself, without using the known definitions, can be considered an abstract model of a part of reality - domain for which the knowledge-based system is created. Part of the project is the validation of the methodology for ontology creation. One of the underlying methodological postulates for designing ontology is a logical procedure from an Upper Ontology through a Core Ontology to a Domain Ontology.

The significant project milestone towards creating of the knowledge base is considered domain ontology. A domain ontology building process could consist of the following stages:

- Design the principles of the ontology creation and its connection to other ontologies.
- Preparing a vocabulary of terms - clarification, sorting terminology of the given domain.
- Creating taxonomy in the given domain, i.e. finding relationships between concepts.
- Definition of classes (types, concepts), suggestion appropriate.
- Adding instances to various ontological types.

When the ontology is built, it is necessary to keep in mind that it is considered domain ontology, therefore the bottom-up or middle-out approaches are expected. When applying the middle-out approach, the ontology is built from the most frequent to the less frequent concepts. Based on the proposed ontology was developed application in JSP programming language that presents the possible type of knowledge-based system.

2 Ontology modeling and design
Following part of paper is situated to ontology engineering area, which is focused to developing and creating ontology.

2.1 Upper ontology selection
In information science, an Upper Ontology (Top-Level Ontology or Foundation Ontology) is an ontology which describes very general concepts that are the same across all knowledge domains. The most important function of an Upper Ontology is to support very broad semantic interoperability between a large numbers of ontologies accessible —under this Upper Ontology. The following ontologies are now competing to be used as the foundation for standard:
- IFF Foundation Ontology.
- Suggested Upper Merged Ontology.
- OpenCyc.
- Lattice of Theories including the above and the 4D ontology based on ISO 15926.
- The Multi-Source Ontology.

2.2 Core ontology meaning and role
In philosophy, a Core Ontology is a basic and minimal ontology consisting only of the minimal concepts required to understand the other concepts. It must be based on a core glossary in a human language, so that humans can comprehend the concepts and distinctions made. Such a Core Ontology is a key pre-requisite to more complete ontology foundation, or a more general philosophical sense of ontology. Core Ontology is a concept that is used in information science as well.

Core Ontology has an important position in the interoperability area. It is a central ontology for systems that integrates many ideas from various points of view of the same problem. Other view on the Core Ontology corresponds with the work of representatives from various communities with the goal of harmonizing their knowledge perspectives [2]. The next solution of the Core Ontology is connected with the integration of dictionaries from many fields of the same theme, for example in medicine, in an attempt to find the core part that is the same (or similar) in all fields [3]. In the MENTAL project the Core Ontology is a general model of the military at the Czech MoD, see Figure 1 and to compare in the doctoral thesis the Core Ontology of the KM area, see Figure 2. These ontologies should integrate all ideas concerning knowledge management in Czech military area.

2.3 Domain Ontology creation
A Domain (or Domain-Specific) Ontology models a specific domain, or a part of the world. It represents particular meanings of terms as they apply to that domain. Since Domain Ontologies represent concepts in very specific and often eclectic ways, they are often incompatible. As systems that rely on Domain Ontologies expand, they often need to merge Domain Ontologies into a more general representation. This presents a challenge to the ontology designer. Different ontologies in the same domain can also arise due to different perceptions of the domain based on cultural background, education, ideology, or because a different representation language was chosen.

The methodology for creating the ontology should include a preparatory stage, in which a set of documents that sufficiently describes a given domain (document base) will be collected. At this stage, the project team members were trained in the fundamentals of ontology, and tried to create a working version of their own
ontology. Furthermore, it is necessary to clarify the basic concepts of the subject area in focus; for instance, by means of the analysis of the document base, which characterizes the selected domain. Basic concepts of the domain are arranged, e.g. into taxonomy.

Taxonomy is a set of concepts, where concepts of higher levels can be further developed by concepts of lower levels. The depth of the hierarchical structure is set by goals for which the ontology is created. Our research team uses the TOVEK\textsuperscript{2} products for this purpose, especially Tovek Tools Analyst Pack. The document base should be put into a unified form, which assumes the selection of documents by language and format. A prerequisite to an appropriate ontology design is a good understanding of the subject area (domain) of the future KMS. This is an iterative — top-down and bottom-up procedure which leads to continuous improvement of the original proposal. The main criterion for the quality of the ontology will be an effective and user-friendly knowledge application.

The ontology design contains a set of ontology classes: area of interest, project, process, document, person, organization, system, equipment and armament, theme, place, capability, stage, rule, procedure and event. Each class has its own definition and attributes that it characterize. The relations between classes are in fact named and described separately, not so important in that place for the purpose of paper. This ontology was implemented in knowledge based application AToM2 in project MENTAL.

The ontology of KM domain, which was used for open-source solution of knowledge-based application in doctoral thesis, has classes: term, process, person, discipline, method, organization, data source, tool, language, technology.

3 Taxonomy versus ontology

Taxonomy is a set of controlled vocabulary terms, which are organized in a hierarchical structure. Each term is included in at least one parent-child relationship. Different types of parent-child relationship can be distinguished, such as whole-part, type-instance, is-a or genus-species. The last one is often used in taxonomy in biology. When building taxonomy, it is useful to set a rule that the parent is of the same type as descendants (this is not applicable to the whole-part type). In taxonomy it is also true that a child has only one parent. If there are more parents, the taxonomies are known as poly-hierarchical. This approach is applied when a concept is found in several places in the taxonomy. If so, then it is understood as the same concept. When building the taxonomy, the number of parents was reduced to one; for building hierarchies, the whole-part or part-instance types were selected. An example of the whole-part relationship is a representation of the organizational structure of departments or institutions; the type-instances relation can be traced in the processing of specific positions and persons in organizations. When building the taxonomy, relations that are not hierarchical were often accessible. These relations are also worth noticing, especially for further processing in the ontology building. A typical example is a simple related-to relationship, the one that is known from building thesauruses.

As already stated, the aim is to create ontologies in the NEC and KM area, a controlled vocabulary which is expressed by means of an ontological language. Ontology thus contains a dictionary of terms that can be adapted from the taxonomy already built; the description of the meaning of the concepts in the context of the given domain, i.e. both among individual concepts and within the context of that domain as a whole [5]. It should be mentioned here that the ontology under construction is supposed to be the basis of the knowledge portal. In this case, the ontology does not necessarily need to be described by a formal apparatus, which ensures the possibility of computer processing and deriving new knowledge. Historically understood, taxonomy is in fact a special type of ontology (a lexical ontology) with a very limited set of rules, and therefore the taxonomy can serve as a knowledge base for the knowledge portal. Within the project, the ontology with a wider set of rules will be built. To ensure the formal correctness, the tools designed for ontologies editing are used, for instance the Protégé using OWL or Ontopia; or ATOM2, the formal apparatus of which is based on Topic Maps technology.

Knowledge-based application is the main subject of our research, because no real-world applications and knowledge operating exists in Czech military area. The purpose is to show the ways how can be application developed.

\textsuperscript{2} TOVEK, 2010, www.tovek.cz
Fig. 1 The MoD Core Ontology

Fig. 2 The KM Core Ontology
4 AToM2 realization

Afer taxonomy and ontology exist, following step is to develop application based on Topic Maps standard. AToM2 system by company AION was supplied like an outsourcing solution, which uses ontology model from Figure 1. This part documents only the current achievements in the knowledge system (KS) construction. The KS that is being built on top of the ontology consists of two parts. The first part is based on common functionality provided by the supporting technology (e.g. ATOM or Tovek). The second part is created based on custom requests. The custom part includes functionality that is not provided by the supporting tool; it includes:

- Personalization.
- Communities of Interest support.
- Bookmarking and history for person, group, document.

The several following images display current stage of KS development based on the taxonomy and ontology. The system is being built for department in MoD of the CR; all views are available only in Czech language; we deeply apologize but we are not able, including the paper writing, transfer the knowledge system content (or part) into English language. The content is too complex and comprehensive result of the several team members.

On the Figure 3, the list of ontology classes and the number of instances is displayed; on the Figure 4, the entry point of the system. The NEC topic on the left (under group label) is the name of the entire ontology. The objects on the right are the main instances, like NEC Domains, NEC Capabilities, etc., of selected classes associated with NEC topic.

On the Figure 5, the example of topic detail is provided. The topic belongs into Document class, the document structuring, the document names and other links based on ontology are available. On the Figure 6 is an example of faceted search. On the first figure the keyword NATO is entered. The list of results is returned (on the right – 20 results). On the next figure, the result is filtered based on ontology classes and groups (all projects connected to search).

Fig. 3 Ontology classes and associations (selection) implemented

Fig. 4 Knowledge System Entry Point
Fig. 5 Example of the topic detail (selection)

Fig. 6 Faceted search – keyword search

Fig. 7. WebAppOnt structure
5 WebAppOnt realization

Other development example of knowledge-based systems has been implemented in the doctoral thesis. The system WebAppOnt is an alternative solution to the project MENTAL particularly to KBS AToM2, which has had devise procedures and methods for developing the knowledge application. The important note is that the system WebAppOnt is a concept developed by one person. It is therefore not elaborate in detail, but this is not a problem. Thanks to the JSP technology based on Java language system can be extended and putting it on the adequate level as a commercial system. The WebAppOnt consists of two basic layers - knowledge layer, which is provided through Ontopia tool and the application layer or efficient layer, which is programmed in JSP environment and its task is to provide required data to the user.

Ontology is the part of the knowledge layer and it is constructed in Ontopia application like an Topic Maps standard. Four main application windows are selected, the ontology window, the browse topic window and searching screen.

On the Figure 8, the list of topic type (ontology class) and its instances are displayed. Instance is described in details with all attributes and associated terms, links and documents from knowledge base. The whole ontology is defined according to the proposed model.

Window for search query is presented on Figure 9:

The result of searching is a list of searched terms related to entered topic. The terms are sorted by relevancy. Relevancy is determined by the occurrence of the term in the selected area which is defined by some data source. The document is part of a knowledge base. Search module of WebbAppOnt application is the subject of Figure 9. Each from the selected terms can be browse through the hyperlinks in background.

6 Conclusion

The article briefly discusses problematic about creating knowledge-based systems. Project MENTAL is still under progress and doctoral thesis is finished. Further details about research and KBS area indicates literature [1][8]. The main acquisitions of article are real applications based on ontology. The applications are part of defense research, but their use is so universal, the need to consider also the real-life usage.

From the project MENTAL view a significant part of the research task was finished and the same part is still to be developed. During the project development new theory, technology and tools were used. This is the first project dealing with the KM theme at the MoD. The aim of this paper was to point out the pitfalls in the ontology development, in the transformation from taxonomy into ontology using an instance of the knowledge-based system in the NEC domain. In retrospect, the taxonomy built from the perspective of the needs of ontology can be subjected to critical evaluation: the taxonomy often contains relationships that are difficult to be transformed into ontology, such as frequent use of the whole-part relationship. Nevertheless, it is obvious that the creation of taxonomy is a vital stage in the process of building ontology.

From the dissertation view was developed and introduced the alternative solution knowledge-based application. This result has excellent value, is free, has a lot of options and more tips to expand it, is able to adapt to the requirements from the users and it is independent on the platform. Other benefits are listed in the doctoral thesis.
Fig. 8. WebAppOnt solution
References:


