THE KNOWLEDGE ENGINEERS PREPARATION

Ladislav Burita
Department of CIS, Faculty of Military Technology, University of Defence, Brno, CR and
Department of IEaIS, Faculty of Management and Economics, Tomas Bata University, Zlin, CR,
Czech Republic
ladislav.burita@unob.cz

Abstract:
The article presents in a case study ideas and methodology of the preparation of knowledge engineers at the university that reflects years of experience in knowledge-based systems development. The mission of the knowledge engineers and the need of their preparation in the information and knowledge society are mentioned. The theoretical base of the knowledge management systems solution is Topic Maps its main parts are defined. The software tool ATOM is described, as an implementation environment of the Topic Maps concept and its modules for the ontology preparation, for the data input, and the user access in portal. The steps of the teaching methodology are mentioned and its key characteristics are explained. The study materials are a part of the knowledge management system. Many pictures from the various projects complement the whole view at the used technology and suggested methodology. The acknowledgement mentions the current project at the University, Faculty and Department.

Keywords: knowledge management system, human capital, knowledge engineer, education, ATOM methodology

INTRODUCTION

The information and knowledge society needs the high quality information systems (IS) and knowledge management systems (KMS), in which will be clearly organized information and knowledge, and these are readily available. The society needs well trained professionals who will be able to work with knowledge in commercial enterprises, organizations and government institutions, and will be able to innovate the business processes. The interest of the European Commission in this matter is declared by means of the aims in the Strategy (i2010, 2005) to develop the information society and its innovations for the next period (i2020, 2010).

Meeting such demands is not an easy task. Therefore, besides an increase in the potential of knowledge, the students should develop creative skills and personality traits that would lead them to the enhancement of their problem solving skills. The issue of knowledge management and creation of knowledge-based systems is urgent and relatively new, still under dynamic development. Some of the knowledge systems have been successfully implemented in practice, as well as the principles of work with knowledge, such as (Knowledge System for Lifelong Learning, 2013). The knowledge systems have a great potential for various areas, for example in university cooperation (MilUNI, 2013), for Network Enabled Capability (NEC) administration (MENTAL, 2011), and also for education.

The term "Knowledge Engineer (KE)" appeared in the 1980s in the first wave of commercialization of artificial intelligence. The purpose of the KE job is to work with a client who wants a KMS created for them or their business. A KE integrates knowledge into computer systems in order to solve complex problems normally requiring a high level of human expertise. Knowledge engineers interpret and organize information on how to make systems decisions (Aylett & Doniat, 2002). The article is oriented in the first part to the description the theoretic base and software tool of the KMS solution; the second part is focused on the explanation of the education methodology. The format of the article is a case study that describes the unique SW ATOM as a tool for the knowledge engineers preparation.

1. TOPIC MAPS - THE THEORETIC BASE OF THE KMS

Why will be the solution of the knowledge system based on the Topic Maps (TM) principle? The main reason is that knowledge representation is intuitive; it keeps information in context, and conforms to human thinking more than other knowledge organization principles.
One of the goals and benefits of the ATOM software (SW), as an implementation environment of the TM principle, is to support the implementation of projects of knowledge systems, especially effective development of powerful web applications. This has necessitated some extensions or treatments of the TM standard.

1.1. ISO Standard 13 250: Topic Maps

The Topic Maps (TM) model (Pepper, 2013) consists of the three basic elements: topic, association between topics, and occurrences of the topics. The TM is standardized in ISO/IEC 13250:2003.
The topic contains denominated subject of interest. It stands for a subject we want to discuss; it is its substitute in the computer. Each topic represents just one subject. It is a place in TM where all known information on the given subject is available by means of relations and occurrences.

Subject is a part of the real world, which is described in TM. Each subject is represented by one topic and it can be anything: a person, thing, entity, process, etc. The subject is unambiguously addressed by identifiers:

- Subject Locator – for information sources.
- Subject Identifier – for other subjects.

Associations represent relationships between topics, are bidirectional, and express relationship between subjects; they are not oriented and may have an arbitrary form:

- Unary – represents the characteristic of a subject, e. g. has-seaside.
- Binary – forms a relation between two subjects, e. g. is-a member of a company.
- N-nary – generic relation, e. g. export (from where, to whom, commodity).

Occurrences are formed by information relevant to a given topic; they can refer to information or they might just contain it. There are two types of occurrences:

1. Internal occurrence that represents the topic characteristics and is placed directly in TM.
2. External occurrence as a reference to an external information source. It connects the information and the knowledge layers.

1.2. Changes in occurrences and associations definitions

A simple set of basic features for TM internal occurrences has been in ATOM software renamed to a Variant type property. The following data types have been further complemented:

- Code and Ident – for the identification of the entity, the uniqueness is checked in the SW.
- Group Tree – a simple built-in taxonomy.
- Selection – forms a one-level code list, applied cardinality 1: N and M: N.
- Text – this feature allows inserting text in XHTML; provides the built-in text editor.
- Picture, File – storage of images and files.

In the definition of associations the following changes were in ATOM finalized:

- Order (Sort) – each association can provide a structure of embedded occurrences.
- Power of relationship in % (Rate), e.g. to express supplier-consumer relationship.
- Hierarchy – a special type of association for Parent-Child relations. This feature is for example the basis for hierarchical view of structure of university.
- The starting types of the ATOM SW associations were binary associations. Unary associations were replaced by an extended set of features; especially Group Tree and Selection. The main reasons for the use of N-nary associations were addressed through Sort, Rate and Hierarchy.

The above mentioned changes in associations are appropriately reflected in the creation of web applications in higher performance of applications and further in saving the lines of code which is needed for the service.

2. ATOM SOFTWARE

The ATOM SW is characterized, its principle of operation described and the details of functional possibilities of the Data Editor and the Schema Editor of the ontology (part of ATOM Studio) are added. The examples and pictures in the Chapter are from the project for military universities cooperation (MilUNI, 2013).

2.1. Basic Characteristics

ATOM is a SW for sharing data with co-workers, customers, or friends via web browsers. ATOM is a non-programming web database SW that does not require special knowledge. Anyone can easily construct a knowledge system on the web. The ATOM web database can be used for intranets with more powerful features than typical shared spreadsheets like SharePoint or box.com. It can be used as a construction kit for building web applications with powerful information retrieval, and for various encyclopaedias, dictionaries, knowledge bases in applications where wiki approaches are not enough. The complete ATOM solution includes three layers (environments):

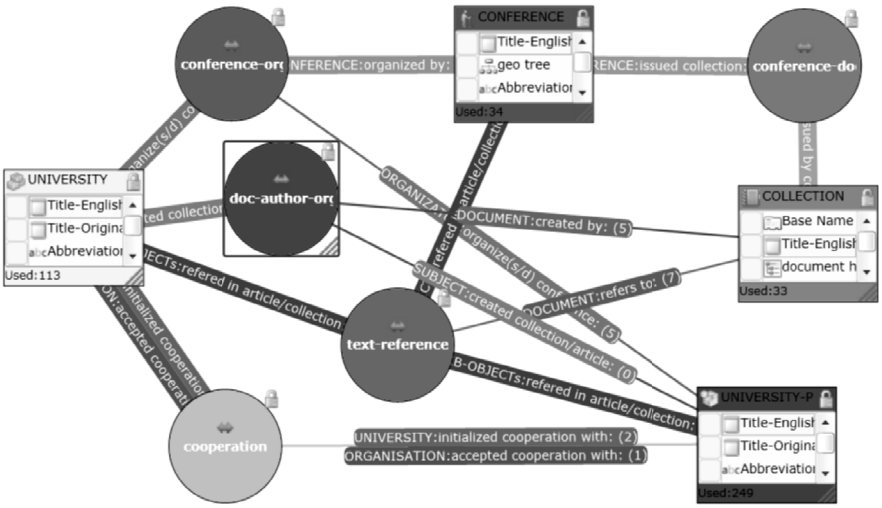
1. ATOM Studio: Ontology Designer, user administration, batch data in/output module.
2. Data Editor: Includes data into database via the ontology.
3. User Portal: Approach to knowledge system.

2.2. How does it work?

At first, create ontology of your problem domain in the Ontology designer module.
You can design the ontology by drawing which is similar to drawing on a flip board (see Picture 1) or writing it through filling in forms and using pre-prepared templates and adjusting them to your needs.

Immediately you can enter data through the forms which are generated from the ontology in the Data Editor module. Data Editor and ATOM Studio are the basic components of the ATOM SW.

Picture 1: Ontology design in ATOM - MilUNI KMS, Schema Editor



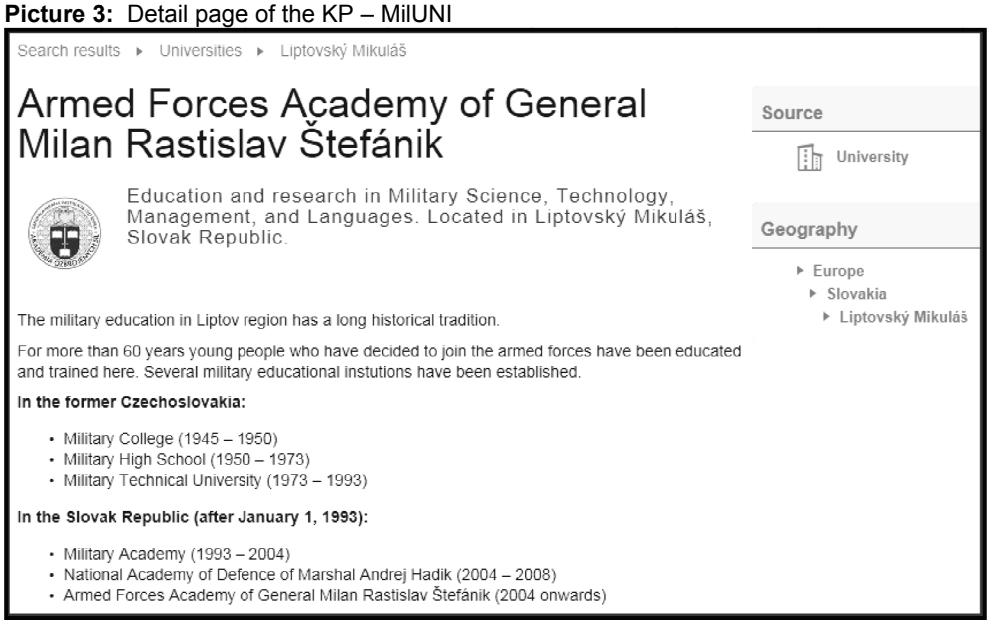
Source: (MilUNI, 2013).

The knowledge portal (KP) covers the knowledge base (KB) in the KMS to shield users from details of implementation. The KP is prepared as a typical portal template that is designed for any similar type of the KMS. The portal structure and functions are designed with respect to the KB content and is connected with the KB through ontology, so that it could be simply edited and personalized. The KP consists of 3 types of pages: Title (see Picture 2), List, and Detail (see Picture 3). The KP includes some menus and boxes for access and search of information. Almost all parts are under administration in ontology.

Picture 2: Title page of the KP – MilUNI

Source: (MilUNI, 2013).

Picture 3: Detail page of the KP – MilUNI



Source: (MilUNI, 2013).

3. METHODOLOGY OF TEACHING THE KNOWLEDGE ENGINEERS

The methodology of teaching is derived from the development of the KMS in projects (MENTAL, 2011), (MilUNI. 2013), and others. The methodology steps, after the theoretic education parts, include:

1. Collecting information sources and their analyse utilizing special SW.
2. Clarification of the terms (ontology classes candidates) and verification them.
3. Ontology design and implementation in Schema Editor of the ATOM SW.
4. Creating and updating the knowledge base.

The task assignment for student work is prepared as a team work; it is intentionally general and ambiguous, so that the students have to search their own approach to the analysis of information sources. For example the domain of the scientific conference:
"Analyse the information sources of the conference, produce an overview of the information systems field, the processing of knowledge, social networks and communication systems that were discussed at conferences".

Without any more details stated in the assignment, the students are expected to carry out analysis of information sources, and to select the articles that cover the given field and examine them in detail. The students are supposed to find the details of each article (about the authors and their workplace, research and implementation tasks carried out, methods and tools used, the results obtained). Consequently, the students are introduced to knowledge approaches, creating ontology and the ATOM SW environment. The assignment is built on the previous activity; for example:
"Create a knowledge base on the conferences; within the knowledge base, process selected articles on information systems, knowledge processing, social networks and communication systems".

3.1. Collecting information sources and their analyse utilizing special SW

The first step is to collect the available data sources. They will be used for searching the basic concepts, candidates for ontology development. In the student work it is a task for teacher to prepare a set of documents as a basis for the student team work.
The set of document is changing. In the article is described an example with the conference domain (task for the students in academic year 2012-13). In this academic year 2013-14 is the theme "Wiki Leaks Documents", one group "Afghan War Logs", other group "Spy Files".

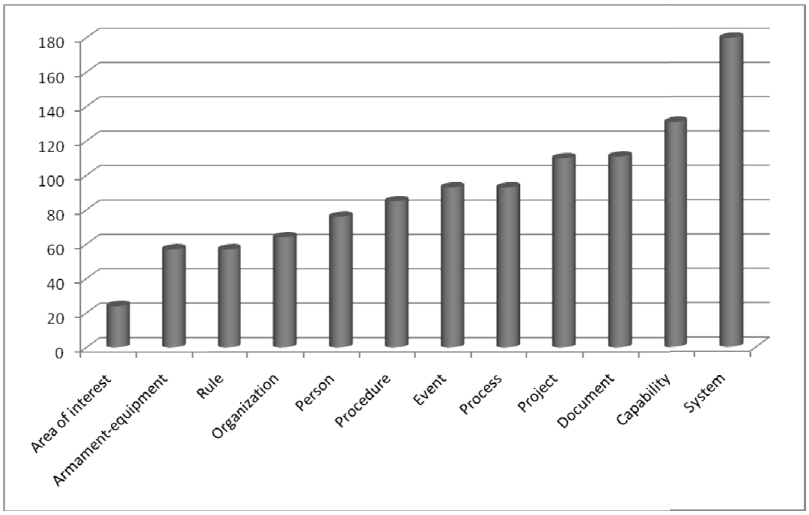
Creating a document base, its analysis and processing is carried out in groups according to the domain orientation of the team members (students) in line with the objectives of the assignment. The

used SW is the Tovek Tools Analytics Pack (TTAP), a professional SW for analysis of information (Tovek Tools, 2013).
The TTAP consist of five modules: Index Manager for indexing of data sources; Tovek Agent for information retrieval; Query Editor for preparation of a complex query; Info Rating for context analysis; Harvester for content analysis. The results of this step are terms for ontology classes' definition.

3.2. Clarification of the terms (ontology classes candidates) and verification them

The candidates for ontology classes' definition are the object for the frequency analysis against the document base; see Picture 4; example from the project (MENTAL, 2011). After that follows the seminar discussion between team members and teacher. Teacher evaluates terms and makes them clear in all aspects and associations.
The result of the discussion is the verified candidates' terms for the ontology classes' definition (conference, proceedings, paper, etc.) that are used in the next step of methodology.

Picture 4: Testing of concepts in document base

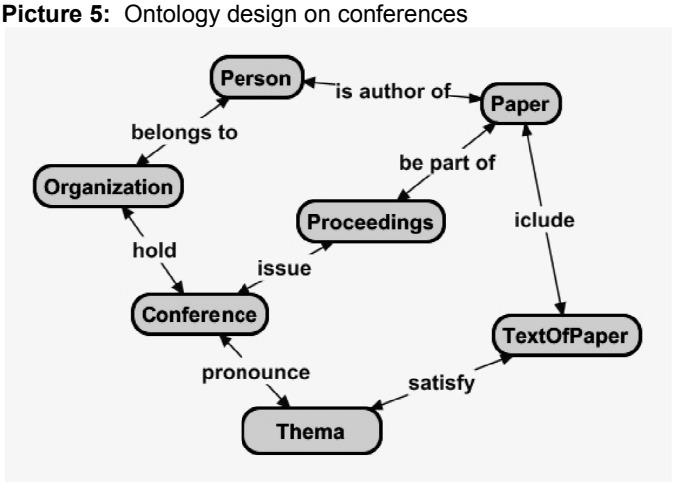


Source: author.

3.3. The ontology design and implementation in Schema Editor of the ATOM SW

The ontology design consists of the two tasks. First are the classes and associations defining in the schema using the SW Visual Understanding Environment (VUE, 2013), see Picture 5. Second are the classes elaborate in the table form and assign characteristics them, see Table 1. Then the ontology is prepared for implementation into the ATOM Schema Editor that is the simple work; see the result at the Picture 6.

Picture 5: Ontology design on conferences



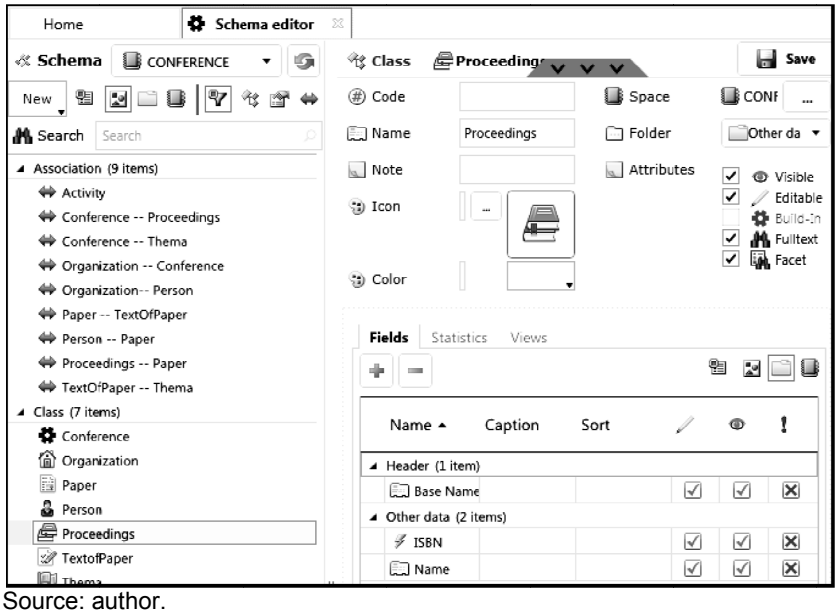
Source: (Burita & Gardavsky, 2013).

Table 1: Classes and their characteristics

	Conference	Proceedings	Paper	TextOfPaper	Theme	Person	Organization
Name	X	X					X
Firstname						X	
Secondname						X	
Title			X	X	X	X	
Abbreviation	X						X
Date	X						
Abstract	X		X				
Place		X					X
ISBN		X					
WWWpage	X	X				X	X
Text				X		X	X

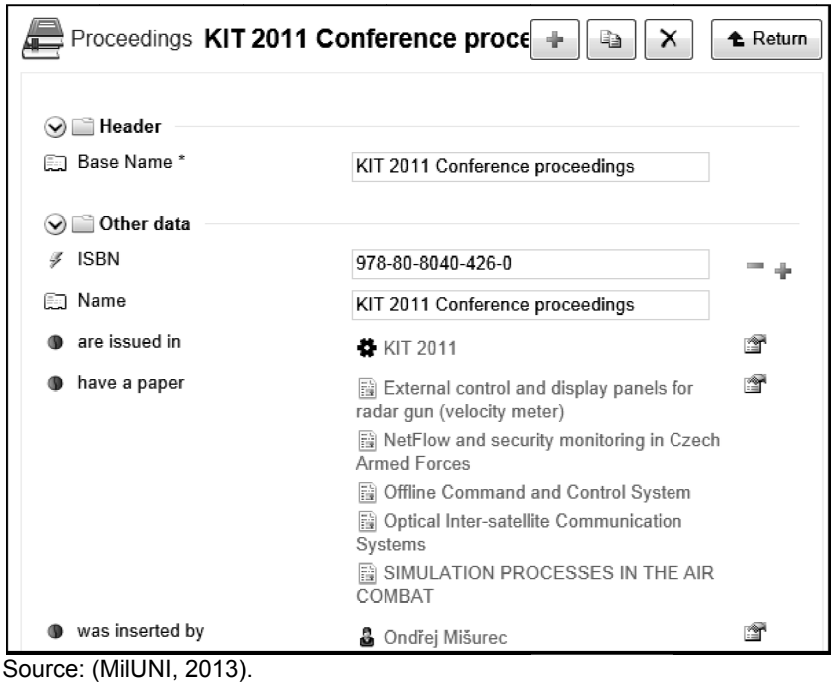
Source: (Burita & Gardavsky, 2013).

Picture 6: Ontology implementation in Schema Editor



Source: author.

Picture 7: Including data in Data Editor

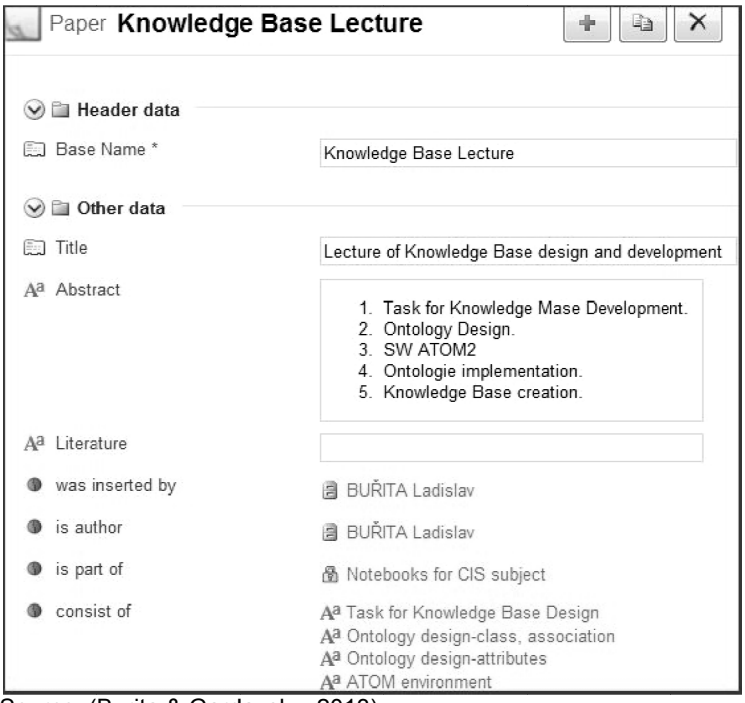


Source: (MilUNI, 2013).

3.4. Creating and updating the knowledge base

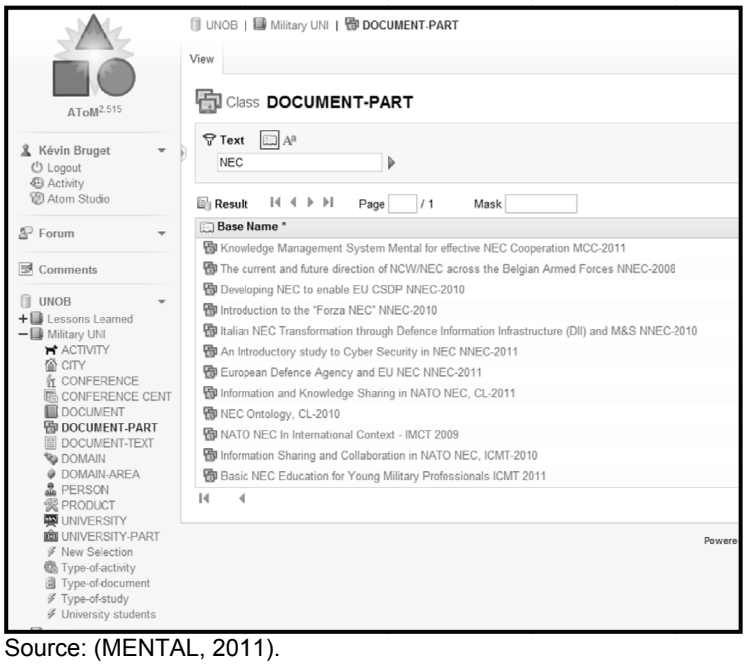
Creating the KB of the KMS is step by step procedure. In the first order are included the occurrences of classes (conferences, proceedings, papers, persons …) and than are the occurrences connected with associations. The data is included in Data Editor, see Picture 7. The result is a complex of information connected together. That information is simple to query, because the conference is connected with the correct proceeding, papers, and authors.
The KMS is not only the source of information, but it can be a source of learning. The opportunity of the KMS for the learning results from its characteristics. The embedded information and knowledge can be divided into small parts and connected in a requirement net, see Picture 8.

Picture 8: Various starting points for the study



Source: (Burita & Gardavsky, 2013).

Picture 9: Search procedure of the paper about NEC



Source: (MENTAL, 2011).

Another possibility to study is using a local class search field and its incorporated filter by putting parts of words or key words in it (example at Picture 9). The next search procedure using classes are based on the summary table which is displayed by clicking on the system name; the user could also reach this menu by clicking on the class. If you know the nature of the searched item or have a global idea of it, this way could be the easiest and the most effective one. Another way is based on the use of the data hierarchy.

4. CONCLUSIONS

The article presents a large variety of areas and ways how the ATOM SW was applied while creating the knowledge-based systems and for supporting learning. The methodology of knowledge engineers preparation is described for the KMS development based on Topic Maps theory. It also introduces the experience and the best method how to teach the creation of knowledge-based systems and how to use them in practice. The created knowledge systems based on the ATOM SW are only at the mere beginning of the application process, but they convincingly demonstrate the wide range of application possibilities.
The SW ATOM is simultaneously the source, tool and target for human capital. The information and knowledge included in SW ATOM could be distributed and used (source of human capital). Education in KMS using SW ATOM leads to the development of creativity and independence of students and prepares them for the work in the information and knowledge society (tool of human capital). The SW ATOM is able to capture the economic, social and cultural complexity of the actual world using its unique features (target of human capital).

5. ACKNOWLEDGEMENT

The article presents the results of more research projects (MENTAL, 2011), (MilUNI, 2013) that were used in the teaching methodology. The current project is (PRO209, 2013) follows the development of KMS, implementation them, and using them in teaching at the University of Defence, Faculty of Military Technology, Department of Communication and Information Systems.

REFERENCE LIST

1. ATOM (2013). Software tool for the KMS development. Zlín, Czech Republic: AION CS. Retrived from www.aion.cz .
2. Aylett, R. & Doniat, Ch. (2002). Supporting the Domain expert in planning domain construction. Retrived from http://planet.hud.ac.uk/aips02kett/aylett.pdf
3. Burita, L. & Gardavsky, P. (2013). Knowledge Systems Applications Based on SW AToM. International Journal of Digital Information and Wireless Communications (IJDIWC) 3 (3)
4. i2010 (2005). European Information Society for growth and employment. Retrived from http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2005:0229:FIN:EN:PDF
5. i2020 (2010). A Digital Agenda for Europe. Retrived from http://ec.europa.eu/digital-agenda/en
6. ISO/IEC 13250:2003. Retrived from http://www.iso.org/iso/catalogue_detail.htm?csnumber=38068
7. Knowledge System for Lifelong Learning. (2013) Retrived from http://www.kslll.net/
8. MENTAL (2011). Project KMS for Network Enabled Capability administration. Brno, Czech Republic: University of Defence, 2008-2011.
9. MilUNI (2013). A portal for military universities cooperation. Brno, Czech Republic: University of Defence. Retrived from http://miluni.eu
10. Pepper, S. (2013) The TAO of Topic Maps. Retrived from http://ontopia.net/topicmaps/materials/tao.html
11. PRO209 (2013). Project for development of organization. Brno, Czech Republic: University of Defence, 2012-2015.
12. Tovek Tools (2013). The analytical software. Retrived from www.tovek.cz
13. VUE (2013). Visual Understanding Environment. Retrived from http://vue.tufts.edu/