Knowledge Management Systems
Based on Topic Maps Theory

L. Buřita¹,², K. Halouzka¹, V. Malý¹* and V. Ondryhal¹

¹ CIS Department, University of Defence, Brno, Czech Republic
² Department of Industrial Engineering and IS, Tomas Bata University in Zlín, Czech Republic

The manuscript was received on 4 December 2016 and was accepted after revision for publication on 16 June 2017.

Abstract:
The aim of the paper is to summarize the experience gained from the research and the constructions of the Knowledge Management Systems (KMS) based on the ontology engineering. Computational ontologies are the means to formally model the structure of a system, to discover the entities and relations that emerge from its observation, and which are useful to the system purposes. In our projects, we utilize the Topic Maps Theory and the tools based on this technology (e.g. ATOM2). The key point of the solution is the ontology development; our methodology procedure is presented; the improvement of the methodology and the future research goals are oriented to automate the procedure steps. During the process, the text mining tools (e.g. TOVEK) and other analytical tools are used. At the end of the paper, the summary of the experience with the methodology application is presented.

Keywords:
knowledge management system, Topic Maps, ontology, ATOM, MENTAL, CEFME

1. Introduction

The paper presents an overview of the authors’ results and experiences in research and development of the Knowledge Management System (KMS). The research started in project „Knowledge Management of the Czech Armed Forces (CAF) Network Enabled Capability (NEC) – MENTAL“ [1].

The aims of the MENTAL were to carry out the analysis of knowledge approaches, ontology and ontology languages, and to assess their suitability for using them inside the CAF; to formalize the ACR NEC strategy and develop an encyclopaedia of NEC terms; to propose a methodology for KMS development; to elaborate the KMS proposal in the CAF NEC administration and to implement it. The accomplishment of the project

* Corresponding author: CIS Department, Faculty of Military Technology, University of Defence, Kounicova 65, 662 10 Brno, Czech Republic. Tel. +420 973 44 35 72. Fax +420 973 44 23 37, E-mail: vlastimil.maly@unob.cz
is assured by successful cooperation of researchers from the University of Defence with the TOVEK [2] and AION [3] companies.

The most important activity concerning the KMS 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.

Why is the solution of the KMS based on the Topic Maps (TM) principle so good? 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. The TM [4] model is described in Chapter 3. The application area of TM is based on the environment of libraries and work with documents, supported by project Ontopia [5].

After the introduction, the literature review follows; the next theoretical chapter focuses on TM as a basis of the KMS; next is the main part of the article about ontology development and KMS examples; the last part is the conclusion.

2. Literature Review

The Knowledge Management (KM) is an important discipline that affects most businesses. Knowledge is now seen as strategically the most important resource and learning as a strategically important capability of enterprises and organizations. Attention to knowledge and its management is growing dramatically and concerns many levels. Without realizing it, the knowledge is issued at every level in any company or institution. Working knowledge is individual and depends on leadership, i.e. how it will be cared for and distributed it among other employees.

Managers questions can relatively easily solved using the knowledge management. Nothing is easier than to ask him whether he knew if the same problem dealt anybody else in the company, how it was solved and what the result was. Many of these manager’s simple questions at least prompts them to reflect the benefits of KM and they begin to wonder about its use, because ultimately it comes to save their time.

The publication of KM: „101 Tips for Leaders in Knowledge-Intensive Organizations“, written by Karl-Eric Sveiby in 1990, started the development of the sector management. Working knowledge simultaneously developed in the Japanese management system. The system, especially its approach to innovation and work with new technologies, was destined to find new ways to manage knowledge.

The KM is described in an article „Identifying and weighing of key performance indicators of knowledge management 2.0 in organizations“ [6]. It is identifying and weighing key performance indicators of KM 2.0 in organizations. According to widespread permeation of technology, especially social media in different organizational dimensions and functional view to this phenomenon in KM, performance measurement of this kind of media in order to meet organizational goals seems necessary. The key performance indicators have been identified and weighed through Delphi methodology and via questionnaire in three rounds.

The article concerning knowledge-based educational environments is „A SKOS-based framework for Subject Ontologies to improve learning experiences“ [7]. The article describes Subject Ontologies that represent conceptualizations of disciplinary domains in which concepts symbolize topics that are relevant for the considered domain and are associated by means of specific relations. Usually, these kinds of lightweight
ontologies are adopted in knowledge-based educational environments to enable semantic organization and search for resources and, in other cases, to support personalization and adaptation features for learning and teaching experiences. For this reason, applying effective management methodologies for Subject Ontologies is a crucial aspect in engineering the environments. In particular, this paper proposes an approach to use SKOS (a Semantic Web-based vocabulary providing a standard way to represent knowledge organization systems) for modelling subject ontologies.

The article above drew attention to knowledge management and its benefit to SW development „10 years of software architecture knowledge management: Practice and future“ [8]. It emphasizes the importance of architectural KM (AKM) for SW development that has been highlighted over the past ten years, where a significant amount of research has been done. Since the first systems using design rationale in the seventies and eighties to the more modern approaches using AKM for design of SW architectures, a variety of models, approaches, and research tools have leveraged the interests of researchers and practitioners. Capturing, sharing, and using AKM has many benefits for SW designers and maintainers, but the cost to capture this relevant knowledge hampers a widespread use by SW companies. However, as the improvements made over the last decade did not boost a wider adoption of AKM approaches, there is a need to identify the successes and shortcomings of current approaches and industry needs from KM.

The connection between KM and SW development can also be found in [9] whose aim is to emphasize the effect of KM practices during SW development projects. This research paper presents the first approach to cope with KM and engineering practices across SW development projects. The main goal is to define a roadmap for representative software development life cycle tasks during a typical software project development. The research introduces an ongoing architectural case study using SW maintenance tasks as a means to enhance the knowledge flows within the organization. Software maintainers validate, correct and update knowledge from previous phases of SW development life cycle through the application of back flushing technique at the data warehouse. Further research developments will present a detailed guidance model for both research areas: KM for SW engineering combining insights across corporate software projects as a means of evaluating the effects on people and organization, technology, workflows and processes.

The book [10] covers, in the 5th chapter of „Next Generation Intelligent Environments: Ambient Adaptive Systems“, key topics in the field of intelligent ambient adaptive systems. It focuses on the results within the framework of the ATRACO (Adaptive and TRusted Ambient eCOlogies) project. The theoretical background, the developed prototypes, and the evaluated results form a fertile ground useful for the broad intelligent environments scientific community, as well as for industrial interest groups. The objective of the knowledge architecture of ATRACO is to enhance communication, as well as to ensure effective knowledge sharing among ATRACO components. It is built around ontologies, ontology managers, and agents. Every component of an activity sphere uses ontology to model its local knowledge and state. These ontologies will certainly be heterogeneous, but they must be used transparently in the context of any sphere. Thus, KM in ATRACO is concerned with the alignment of heterogeneous ontologies, in order to produce the Sphere Ontology, which encodes the sphere knowledge. In this chapter, we shall describe the ontology management framework developed in the context of ATRACO.

The close association between KM and TM is described in [11] and it defines how managers can organize pieces of unstructured information making it semi-structured in
order to ease searching and sharing. Author discusses a new technology called topic maps for information management. Author suggests the use of managerial notes based on TM standard. The timeliness of this subject is well illustrated by the war among the large SW vendors like Microsoft, IBM, Novell, and Google for the potential customers with demands for technologies handling unstructured information.

Research is showing that TM together with the Resource Description Framework (RDF) can provide a foundation for the Semantic Web [12]. They can serve to represent information currently stored as database schemas (relational and object). Where databases only capture the relations between information objects, topic maps also allow these objects to be connected to the various places where they occur.

The close association between KM and TM is explained in [13] and it examines what role modern IT tools play in the process of knowledge sharing inside the enterprise. Author’s hypothesis is that semantic data models and semantic web applications foster knowledge sharing through the codification of tacit knowledge. Authors propose a TM based technology for the management of unstructured information. The technology enables visualization of concepts through associations similar to human thinking and unlike the traditional hierarchical approach; it requires the clear definition of all topics to ensure that all co-workers share a common meaning of each topic and it provides the possibility to attach documents to occurrences of a topic.

The TM in connection with KM is presented in [14] and it explains the distributed heterogeneous engineering KM approach for the practice of collaborative product design. Authors firstly propose to describe an XTM based heterogeneous knowledge integration model in collaborative product design. Then, a knowledge management framework in collaborative design process is developed. Finally, a knowledge management system in collaborative design is constructed.

The KM principles and techniques should be applied to manage SW testing knowledge. A mapping study is described in [15] giving a survey on existing research on KM initiatives in software testing, in order to identify the state of the art in the area, as well as to suggest the future research. Aspects such as purposes, types of knowledge, technologies and research types are investigated.

3. The Topic Maps Theory

Topic maps technology is the ISO/IEC 13250:2003 standard, latest revision was published as ISO/IEC 13250-2:2006. The standard consists of nine parts; the individual parts provide overview, data model, formal model and different syntaxes used for persistence and information exchange like XTM (XML Topic Maps), CXTM (Canonicalization XTM), CTM (Compact TM Notation), GTM (Graphical TM Notation).

3.1. Informal Approach

Informally, topic map consists of three main elements – topics, occurrences and associations. A topic, in general, can be any „thing“, a person, an entity or a concept, anything about which may be asserted by any means whatsoever. Subject is defined as the real world „thing“ that the topic itself stands in for. Topic refers to the object or node in the topic map that represents the subject being referred to. Topics can be referred to by their type; it is the way we can categorize them. The relationship between a topic and its type is an example of the class-instance relationship. Topic may be linked to one or more information resources, relevant to the topic. Such information resources are defined as the topic occurrences. Occurrence may be any resource from books about the topic to
image, video covering the topic to short comment in discussion about the topic. Occurrences are usually external to topic map, depending on used syntax we can point to resource using URI (Unified Resource Identifier) or by any other means. Occurrence can be specified by its role and role type. The last fundamental part of the topic map is the topic association. The association declares relationship between two or more topics. Association can have a type, in the same way as the topic or occurrence. In addition, each topic plays role in the association. Roles can be typed as well.

3.2. Formal Model

Formally, topic maps can be described using proxies. The model is based on ISO/IEC 13250-5 (part 5) Topic Maps Reference Model.

According to this reference model, there is only one representative for subjects we described in informal description; subject proxies (proxies). The topic map (subject map) is then an infinite set of proxies. The set of all such maps is denoted as $M$.

![Proxy structure](image)

Fig. 1 Proxy structure

A proxy is a finite set of properties $\{p_1, p_2, ..., p_n\}$, with $p_i \in P$. A property is a key/value pair $<k, v> \in L \times V$, where $L$ is finite set of (proxy) labels, $V$ is a finite set of (proxy) values; numbers, string, etc. Property may contain recursively references to other proxies. According to the definition, a particular book can be represented by the following proxy

$$\{(\text{name, "A Brief History of Time"}, (\text{pages, 212}), (\text{isbn, 0553380163})\}.$$

3.3. Comparable Technologies

There is another widespread technology used as a background layer for knowledge management systems construction. Resource Description Framework (RDF) is a language for Semantic Web, the language for description of the resources or the metadata on the Web. The comparison of the approaches can be found in [16]. The main difference between both approaches is that topic maps focus on the knowledge representation from the point of view of a person, a human being; meanwhile RDF focuses on the support of machine communication and reasoning.
4. Ontology Development

The methodology of an ontology development we have suggested in the research project MENTAL, described in [17] and verified in the project MilUNI [18]. The methodology steps include:

1. Collecting the information sources and their analyses by utilizing special SW.
2. Clarification of the topics (candidates of the classes) and their verification.
3. Designing the ontology and its implementation in ATOM Schema editor.

The methodology steps are explained in an example „Network Enabled Capability (NEC)” that involves the analysis of the domain of NEC in the defence research project MENTAL. The goal of the project was analysis, design, and implementation KMS for administration of the NEC concept in the Czech Armed Forces (CAF).

4.1. Collecting Information Sources and their Analysis

The first step is to collect the available data sources, then to create the document base and further to generate the best understanding of the domain of interest; next, to specify the proper set of analysed themes, and finally to choose the candidates to classes for the ontology development.

The used software (SW) is TOVEK, a professional SW for analysis of information sources and text mining. SW TOVEK consists of five modules:

- **Index Manager** for indexing data from miscellaneous information sources (files, web pages, E-mail, databases). The indexed resources allow users quicker access to information;
- **Tovek Agent** is the interface for searching in indexed documents from local or from Tovek Server information sources and their logical presentation. The result of the retrieval (hit) can be accessed using the Tovek viewer;
- **Query Editor** for preparation a complex hierarchically structured query. The search of document can be affected by the using of individual query parameters and weight;
- **Info Rating** for contextual analysis that allows the user to easily sort and search documents by selected topics. The result of the module is in the contextual matrix or graph; and
- **Harvester** for content analysis that allows the user to detect keywords in the documents and their connections to each other.

The other analytical or text mining SW could be used, of course. The source for analytical text process includes (as in the beginning of year 2016) more than 300 documents.

4.2. Clarification of the Topics (Classes) Candidates

The candidates for topics (classes) of the ontology definition are an object for the frequency analysis (quantity of occurrences in the documents) against the document base, using the TOVEK SW; see Fig. 2

The next tool used for understanding the domain is a taxonomy, see Fig. 3; the used SW is Free Mind [19]. The result of the phase is a good understanding of the domain; verification of the classes candidates for the ontology definition. The next method used for verification is a team discussion – brainstorming.
4.3. Designing Ontology and its Implementation

The ontology design consists of two steps. The first step is to suggest a schema of classes and associations between them, depicted in the SW VUE [20], see Fig. 4.

The set of classes consists of the following: AREA of INTEREST (AOI), CAPABILITY (CAP), DOCUMENT (DOC), EQUIPMENT (EQU), EVENT (EVE), ORGANIZATION (ORG), PERSON (PER), PLACE (PLA), PROCEDURE (PCD), PROCESS (PCS), PROJECT (PJ), RULE (RUL), STAGE (STA), SYSTEM (SYS), THEME (THE).
The Fig. 4 shows the ontology classes and associations among them. The associations are numbered and their captions and meaning is expressed in the table that is not a part of the paper (it is too long).

For example, association number 29 illustrates that the text of document contains any area of interest; association number 5 shows that the document is about any project; association number 9 suggests that any system was developed by any organization; and association number 32 shows that any person participated in any process, etc.

The second step is the table preparation with attributes assignments to classes, see Tab. 1. Then the ontology is prepared for the implementation in the ATOM Schema Editor; see the result in Fig. 5.

Tab. 1 Ontology design on KMS, classes and attributes

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Class:</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Name</td>
</tr>
<tr>
<td>1</td>
<td>Name</td>
</tr>
<tr>
<td>2</td>
<td>Code</td>
</tr>
<tr>
<td>3</td>
<td>Title</td>
</tr>
<tr>
<td>4</td>
<td>Address</td>
</tr>
<tr>
<td>5</td>
<td>Annotation</td>
</tr>
<tr>
<td>6</td>
<td>Evaluation</td>
</tr>
<tr>
<td>7</td>
<td>Risk</td>
</tr>
<tr>
<td>8</td>
<td>City</td>
</tr>
<tr>
<td>9</td>
<td>TechData</td>
</tr>
<tr>
<td>10</td>
<td>Employer</td>
</tr>
<tr>
<td>11</td>
<td>DateStart</td>
</tr>
<tr>
<td>12</td>
<td>DateFinish</td>
</tr>
<tr>
<td>13</td>
<td>Email</td>
</tr>
<tr>
<td>14</td>
<td>WWW</td>
</tr>
<tr>
<td>15</td>
<td>Figure</td>
</tr>
<tr>
<td>16</td>
<td>File</td>
</tr>
</tbody>
</table>
Knowledge Management Systems Based on Topic Maps Theory

5. KMS Examples

The knowledge management systems (KMS) development have started in defence research project MENTAL for Network Enabled Capability (NEC) administration of the Czech Armed Forces (CAF). The project result has been implemented in Staff IS and it is available for all CAF members.

The successor of this project was MilUNI project [16] for cooperation support of military universities and its result is available at http://www.atom.miluni.eu/. The MilUNI portal was a source for derivation solution to support the CEFME (Central Europe Forum for Military Education) [21] community and it is described in Chapter 5.1. CEFME portal is the most successful KMS system solution of the team and it is accessible on http://beta.cefme.eu/.

The KMS development is connected with a relatively huge „manual“ work with data. The research activities of the team in „Lessons learned from the military missions“ [22] project were oriented to the automation data procedures.

5.1. The CEFME portal

The goal of the CEFME (Central European Forum on Military Education) group (see Tab. 2) is to support education and training of military professionals and cooperation of the participating universities in teaching, research, exchange of teachers and students to increase mutual awareness and facilitate the organization of joint events.
Tab. 2 Universities of the CEFME group

<table>
<thead>
<tr>
<th>State</th>
<th>University</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Czech Rep.</td>
<td>University of Defence</td>
<td>Brno</td>
</tr>
<tr>
<td>2 Estonia</td>
<td>Baltic Defence College</td>
<td>Tartu</td>
</tr>
<tr>
<td>3 Croatia</td>
<td>Croatian Defence Academy</td>
<td>Zagreb</td>
</tr>
<tr>
<td>4 Hungary</td>
<td>National University of Public Service</td>
<td>Budapest</td>
</tr>
<tr>
<td>5 Poland</td>
<td>National Defence University of Warsaw</td>
<td>Warsaw</td>
</tr>
<tr>
<td>6 Austria</td>
<td>Austrian National Defence Academy</td>
<td>Vienna</td>
</tr>
<tr>
<td>7 Romania</td>
<td>Carol I National Defence University</td>
<td>Bucharest</td>
</tr>
<tr>
<td>8 Slovakia</td>
<td>Armed Forces Academy of Gen. M.R. Stefanik</td>
<td>Lipt. Mikuláš</td>
</tr>
<tr>
<td>9 Slovenia</td>
<td>Military Schools Centre</td>
<td>Maribor</td>
</tr>
<tr>
<td>10 Serbia</td>
<td>Military Academy Beograd</td>
<td>Beograd</td>
</tr>
</tbody>
</table>

The portal is suggested not only for the cooperation among CEFME community members, but it has been excellently utilized for the cooperation with other relevant communities; for example LoD7 (working group established in 2009 by the Secretariat of European Security and Defence College to control the international project „Exchanges of young officers, inspired by Erasmus”) and iMAF (International Military Academic Forum).

The CEFME portal shares information about CEFME universities and cooperating iMAF and LoD7 communities. All information concerning universities is collected, provided and shared by these classes: university, university-part, organization, product, geo-tree (counties and cities), person, conference, collection (journal, conference proceedings, or other form of set of articles), article, venue, activity (program, project, and product), and domain tree (areas of interests). The new classes created for CEFME are as follows: meeting, action, workgroup, and function. The objects' information is linked to the CIA FactBook [23] in order to use information about countries and cities.

It is an internet-accessed application; its processing is fully outsourced by SWATOM developer Company AION CS, Zlín, with excellent cooperation experience. The structure of the portal is based on knowledge principle. It means that the information is put in context and it builds the network; it is possible to browse in one query from one piece of information to another without interruption. The system contains a very powerful full text search engine and it integrates the information about objects and with the GIS information. There are 3 frameworks in the system:

1. **Portal** for user access.
2. **ATOM Studio** for ontology definition and administration of users, see Fig. 5.
3. **Data Editor** for data input and update, see Fig. 6.

The portal consists of the 3 types of pages:

1. The title (main) page, see Fig. 7.
2. The search result list, see Fig. 8 (set of conferences).
3. The detail of object information, see Fig. 9 (conference).
Fig. 6 Data editor – set of conferences

Fig. 7 Homepage of portal
The homepage of the portal displays news and offers the access menu with special boxes. The menu structure and content is defined in the ontology and it can be easily changed. The search result list offers a set of objects to choose from as a result of the search procedure. The detailed object information displays the information in a detailed form from the selected items.
5.2. Portal Administration

The portal adoption and use depends on meeting the requirements of users. One of the requirements is the timeliness of information which can only be achieved by involving national administrators and their active participation at own university. The first step towards this activity was achieved by completing training at the premises of University of Defence, attended by nine universities, mostly by IT specialists or staff of Foreign Relations Departments.

The administrator should act at its own university as an initiator and stakeholder in working with the portal in the CEFME community. The research team is trying to simplify its task as much as possible. In the research team a person, whose responsibility is to communicate with the national administrators, was assigned. The cooperation must be prepared in such a way that every problem of the national administrator in portal data input or edit is solved as soon as possible.

Supporting elements as the structure and the content of the portal are integrated for this reasons and it will be further expanded. First, there is a list of addresses of contact persons and national administrators, as well as contacts to the research team and supplier organization. Furthermore, there is an overview of the underlying changes in the portal, FAQ (Frequently Asked Questions), set of predicted queries to the portal with the appropriate answers. In the portal, a range of sample model objects for editing own data can be used. What is still lacking is the context-sensitive help, but software ATOM contractor is working on it.

To communicate with the national administrators, the „mailing list“ can be used which is based on the registration that ensures automatic messaging. One can easily imagine that the content of the report may be a requirement for an action of administrator, or an alert to changes in the structure of the portal, or rules of inserting or updating data. All the above can be facilitated by the distance administration of the portal and an effective cooperation between national administrators in the maintenance and improvement of the CEFME portal.

6. Conclusion

The aim of the paper is to summarize the experience gained from the research of the KMS and it describes the possibilities for large-scale use of KM in education. The article presents relevant basic terms concerning the theoretical part of The Topic Maps, as well as it presents practical basis of the knowledge approach in information processing that is based on SW ATOM.

The key point of the solution is the ontology development, which was designed and realized in the defence research project MENTAL and it was successfully verified in military universities education and training project MilUNI. The methodology steps include three steps relating to collecting the information sources, clarification of the topics and their verification, and finally the ontology designing and implementation in ATOM Schema editor. The improvement of the methodology and the future research goals are oriented to automate some steps.

The consequent research could be oriented to the comparison with similar KMS solutions in military environment, especially in the theme of the NEC [24]. It might be interesting to compare goals, approaches, tools, models, results, etc.
Acknowledgement

The article presents the results of the research activities in projects [1, 18, 25], containing analysis, development, implementation KMS, and using them in CAF and teaching at the University of Defence, Faculty of Military Technology, Department of Communication and Information Systems.

References


