Developing knowledge management systems: approaches, technologies and methods

Bostjan Grasic, Vili Podgorelec
University of Maribor
Faculty of electrical engineering and computer science
Institute of informatics
Smetanova ulica 17, 2000 Maribor
Slovenia
{bostjan.grasic, vili.podgorelec}@uni-mb.si

Abstract: Knowledge management is becoming one of key success factors for organizations in global economic environment. Knowledge management in the scope of computer science is widely understood as a process that supports and facilitates knowledge creation, access, modification and reuse. In the scope of this paper we investigate different approaches for deploying knowledge management. We are especially oriented towards mid-sized organizations with limited budget and limited resources with knowledge engineering capabilities. We propose an ontology based knowledge management system and present the core architecture as well as a development method for such a system.

Key–Words: knowledge management, knowledge management system, knowledge base, ontology, software engineering

1 Introduction

Knowledge is one of key competitiveness drivers in modern economy. In order for organizations to be able to expand in global economy, they have to find value added in them selves, they have to innovate and use the knowledge that lies in their employees. Knowledge management deals with issues how to manage knowledge inside organizations in order to empower innovation, as well as quality and efficiency. There are several definitions of knowledge management, one of them is given by Brian Newman, who defines knowledge management as the collection of processes that govern the creation, dissemination, and utilization of knowledge. Knowledge management by itself is not a technical discipline, however, computer science deals with technical aspects of knowledge management in order to support the knowledge management process.

In this manner various systems and approaches for knowledge management have been researched and developed. In this paper we investigate various existing knowledge management systems and approaches that are suitable for organizations that endeavour to incorporate knowledge management in their business processes. We describe deeply ontology based knowledge management solutions as well as methods for developing ontologies and ontology based systems.

The organization of the paper is as follows. Next section provides an overview of knowledge management systems. In Section 3 ontology based knowledge management is described in detail. Section 4 provides an overview of ontology development methods and introduces a new method for developing ontology based knowledge management systems. The final section concludes the paper.

2 Knowledge management systems

In the scope of computer science, academia has widely agreed upon the role of knowledge management systems (KMS). Knowledge management systems provide IT support for knowledge management process. Widely accepted definition of knowledge management in the scope of computer science is provided by O’Leary, who defines knowledge management as formal management of knowledge for facilitating creation, access, and reuse of knowledge, typically using advanced technology[7]. Benjamins et al. [1] specify the role of KMS by specifying actions, which the knowledge management process should involve. They have defined following actions [1]:

- Knowledge gathering: acquisition and collection of the knowledge to be managed.
- Knowledge organization and structuring: imposing a structure on the knowledge acquired in order to manage it effectively.
Knowledge refinement: correcting, updating, adding, deleting knowledge (maintaining knowledge).

Knowledge distribution: bringing the knowledge to the professionals who need it.

In this manner, an arbitrary KMS should provide IT support for the knowledge management process, by providing tools for the actions specified above. However, looking from the practical point of view, even the best system is worthless if it is not accepted by end users. Based on our experience, we have provided a set of requirements for knowledge organization inside a KMS. In order to empower utilization of KMS, knowledge in a KMS should be:

- unified - everybody agrees to the information specified in a KMS,
- carefully prepared and validated - knowledge is properly structured and is valid in the given domain,
- systematically maintained - there is no ambiguity in the information provided by KMS,
- machine readable - computers are able to process and reason on information provided by KMS,
- reusable - same pieces of information may be reused in different, unrelated business processes and information systems.

When deploying knowledge management systems there are two options. Organization may deploy an existing solution or build a custom-tailored solution. In the rest of this section we are describing different options that are available for organizations seeking to deploy KMS. We have divided KMS into two subsets: common knowledge management systems and intelligent knowledge management systems. The difference between those two categories is in the way knowledge is represented. Knowledge in common KMS is not machine readable, therefore computers are not able to process knowledge automatically. In intelligent KMS, knowledge is represented in a machine readable manner. Because of this, computers are able to have an awareness of the information stored in a KMS. Therefore, computers are able to process information automatically and deduce new knowledge from the knowledge stored in the KMS.

2.1 Common knowledge management solutions

There are plenty of common knowledge management solutions available on the market. We are describing most significant categories of common KMS. For each category there are a number of proprietary, freeware or even open-source solutions. When considering a common KMS, a special care should be taken on choosing a system with appropriate knowledge granularity. Depending on the system, the core information unit may be a sentence, an article, a topic or even a whole file.

2.1.1 Content Management Systems (CMS)

Content management systems (CMS) allow a group of users management of work flow in a collaborative environment. There is a number of different CMS available, most of them are web based. A common characteristic of all CMS is: the higher the number of active users, the greater is the benefit of the system. Granularity of knowledge varies from solution to solution. Generally, the core information unit is a free text article that may be accompanied with multimedia (e.g. photographs, video clips, links to other resources). Common characteristics of CMS are: (i) they enable publishing and exchange of information between a larger group of people, (ii) they aid in easy storage and retrieval of data and (ii) control access to data based on user roles.

The structure of the contributions may be prescribed either loose or strict. Loose structure enables greater freedom at designing and representing the content, while a strict structure on the other hand, specifies a rather precise structure of the content. Weakness of the strict structure is reduced flexibility, while a loose structure often leads to diverse content, which may impose greater problems in the long run.

2.1.2 Document management systems

Document management systems are well established in organizations already. They are mostly used in a role of document repositories, however they may also be used for knowledge management purposes. In this case, the core information unit is a document. They are not particularly suitable for knowledge management since knowledge is distributed between documents. In order to be able to use knowledge, one has first to find the right document, then examine the document and find the needed information inside a document. This procedure is not user friendly and it is very hard, if not impossible to automate it.

2.1.3 Wiki

Wikis are widely spread and accepted on the Internet. Wikipedia is the most well known wiki system and is being one of the most used resources on the Internet for several years. Wikis are web portals that
enable publishing and editing of interconnected web pages. They are mostly encyclopedically organized, meaning each page deals only with one topic. Inside each page there are usually hyperlinks to other topics. Wikis may be either public or private, while contributions may be unrestricted, restricted, moderated or unmoderated. Although they are widely accepted and well understood to general public there are still some weaknesses. Contribution is mainly done using a simplified markup language which may not be appropriate for less skilled users. The main weakness however remains in knowledge representation, which is not machine readable. Computers are not able to process information automatically in wikis, therefore they are not able to provide answers to even simple questions like: which countries lie in EU.

2.1.4 Semantic-Wiki

Semantic wikis are bridging the gap between common and intelligent KMS. They are extension of Wikis described above and are even equally structured as well as very similar looking to common Wikis. The main difference between them is that semantic wikis add machine readable bits of information to ordinary wiki pages in form of semantic annotations. Semantic annotations enable computers to have an awareness of the content inside wiki pages. In this manner, computers are able to process information and provide answers to more complex, content-related questions (like example above). While semantic annotations are very useful on one hand, it is fairly complex task to specify semantic annotation inside a wiki system properly. There are no strict rules for specifying metadata, therefore semantic annotation may become very unsystematic.

2.2 Intelligent knowledge management systems

The main advantage of intelligent knowledge management systems is that knowledge is machine readable. In this manner, computers are able to process information, answer to more complex questions and deduce new knowledge. Because knowledge has to be represented in a machine readable manner, implementation of intelligent KMS is significantly more complex than implementation of common KMS. Knowledge representation is highly dependant on the application domain as well as on specific requirements of the KMS. Therefore it is not expected to use a specific intelligent KMS in different or even similar domains. It is often even a problem reusing the same solution in the same application domain under similar specific requirements. Because of the given facts, there are very few "off-the-shelf" intelligent KMS. Most of them are custom made for a given knowledge management problem. In the rest of this section, we describe briefly common approaches for developing intelligent KMS.

2.2.1 Expert systems

Expert system is a special kind of information system that holds knowledge about a specific domain and is capable of reasoning in that domain. The main objective of expert systems is problem solving in the given domain (e.g. loan approval). Expert systems are composed of knowledge base, reasoning engine and a user interface. Knowledge base is holding a machine readable knowledge representation of the given domain. The reasoning engine is responsible for inference on the facts stored in knowledge base. End users interact with the expert system through a user interface.

Most available expert system are highly specialized to a narrow domain, therefore they are mostly not suitable for adoption as a general knowledge management system. However, there exist also expert system shells, which enable more straightforward development of expert systems. The weakness of expert system shells is that shells usually offer only a framework with the inference engine. The main challenge while developing expert systems is appropriate knowledge representation of the given domain. In this manner, expert system shells are not more suitable for developing knowledge management systems than custom build solutions. Expert systems are only suitable in specific areas, e.g. when an information system should behave similar to a human expert solving a specific problem.

2.2.2 Logic programming

As we have already pointed out, machine readable knowledge representation is common to all intelligent KMS. Knowledge representation is part of artificial intelligence and is dealing with creating models of reality in such a manner that computers are able to reason about reality in a similar manner as humans do [4]. Logic is widely used for knowledge representation (e.g. F-logic, description logic, etc.). In this manner, logic programming may be defined as use of logic for computer programming. There exist several logic languages, prolog and lisp being most widespread.

Logic programming is very suitable for machine readable knowledge representation, however knowledge should be very carefully structured by specially trained specialists - knowledge engineers. Developers that are not used to logic programming, may find hard to combine logic programming with other conventional programming techniques (e.g. object-oriented
programming). They may even have problems switching to the logic programming paradigm. Logic programming is similar to expert systems in that sense that there exist no readily available logic programming based solutions for general knowledge management systems. Logic programming based KMS are generally tightly coupled to a narrow domain, therefore such systems are commonly custom made. Problems may arise while trying to integrate logic programming based systems into information infrastructure because of fairly poor interoperability between logic programming languages and modern programming languages and technologies.

3 Ontology based knowledge management systems

Ontology based knowledge management systems (OBKMS) belong in the category of intelligent KMS. They use ontologies for machine readable knowledge representation. Ontology describes the subject domain using notions of concepts, instances, attributes, relations and axioms. Gruber provided a widely accepted definition of ontology that defines ontology as a formal and explicit specification of a shared conceptualization [5]. Each of the terms used in the definition should be interpreted as follows [8]:

- conceptualization refers to an abstract model of a phenomenon in the world based on concepts identifying this phenomenon,
- explicit means that the types of concepts and their restrictions are explicitly defined,
- formal refers to the fact that the ontology should be unambiguous and machine readable,
- shared reflects the viewpoint that the ontology should cover general knowledge of a domain - knowledge that is not recognized by a single individual, but it is agreed upon a group of users.

Comparing OBKMS to other intelligent KMS there is a similarity in availability of out-of-the-box products. To our best knowledge, there are no domain independent solutions available for general knowledge management. However, there is a rich support for development of ontology based solutions. There exist many tools, guidelines, methods and methodologies for developing ontologies, as well as technology for implementing ontologies and providing interoperability between ontologies and common programming languages. In this manner ontology based KMS are more suitable than other earlier described intelligent KMS.

3.1 Ontologies and Semantic Technologies

Ontologies are one of the key technologies in the evolving Semantic Web. There exist several languages as well as several formalisms to capture knowledge and represent it in ontologies. According to the most recent survey that analyzed the use of ontological languages, OWL is being by far the most used one. Over 75 percent of respondents are using OWL [3].

Web ontology language (OWL) [6] is based on description logic and is one of the main building blocks of the Semantic Web technologies (SWT). SWT is a set of technologies, tools and recommendations proposed by World Wide Web consortium (W3C), that follow the vision of semantic web. The vision of semantic web is in evolution from the web of documents (as we know the world wide web today) into the semantic web (SW). In SW, computers will have an awareness of the meaning of data; hence computer agents will be able to find and process information based on their meaning [2].

SWT provide technologies to achieve this vision. They provide standardized means to represent knowledge on the web scale in such a manner, that computers will be able to process and reason upon it. Main SWT are: (i) RDF (Resource Description Format) - a XML-based language for describing resources on the Web, (ii) OWL - ontology language which is compatible with RDF, (iii) SWRL and RIF - languages for specifying and interchanging rules. SWT provide mature technology for machine readable knowledge representation which is compatible with modern software development paradigms, modern programming languages and Web-based applications. In this manner SWT is a reasonable choice for development of ontology based KMS.
3.2 OBKMS architecture

In order to achieve greater interoperability and flexibility, OBKMS should be split into three layers: knowledge base, reasoning engine and user interface. The proposed architecture of such system is shown in Figure 1. The core component is knowledge base. It consists of the ontology that holds machine readable knowledge of the domain, and of the access interface that exposes functionality of the knowledge base to the user interface, as well as to other applications. The access interface itself may be realized by an arbitrary technology, however we recommend using services in a service oriented architecture manner. This way, the knowledge base itself is truly interoperable with legacy applications. Above the knowledge base is the reasoning component that uses rules to infer new knowledge from the one available in the knowledge base. The user interface should provide users all the functionality expected from a knowledge management system (knowledge creation, access, modification and reuse). The user interface itself may be realized as a single application or a group of specialized complementary applications.

4 Methods for developing ontologies and ontology based systems

There exist several methodologies for building ontologies. We have divided most relevant ones into following groups: early methodologies, methodologies of intensive research and mature methodologies. Early methodologies usually address only one aspect of ontology development process. They are dealing mainly with knowledge representation challenges for building complex artificial intelligence (AI) systems. Generally, they are not dealing with ontology life cycle.

Methodologies of intensive research are building upon experiences learned from early methodologies. They are introducing life cycle to ontology development and are mainly oriented towards development of large isolated ontologies. These methodologies are usually application oriented.

Mature ontologies are reusing concepts introduced by early methodologies and methodologies of intense research. They address ontology development in a holistic manner. Their main contribution is in reuse, collaboration, ontology evolution, ontology evaluation and networking/interlinking of ontologies. Most relevant methodologies for all groups are shown in Figure 2. Beneath each methodology it stands what concepts has particular methodology introduced. The arrows between methodologies indicate reuse of concepts and techniques between methodologies.

4.1 Proposed method for developing KMS

We have developed a method for developing ontology based knowledge management systems that is oriented towards mid-sized organizations with limited budget and limited resources with knowledge engineering capabilities. Our work is based on methodologies described in previous Section, mostly on Onto-knowledge [9] and NeOn methodologies.

The proposed method is based on 6-phase iterative ontology life cycle model. Based on the architecture described in Section 3, this method divides activities into three categories: development of the ontology for knowledge representation, development of access interface to the knowledge base and development of applications that enable creation, modification and access to knowledge stored in a knowledge base. Figure 3 shows the development model as well as activities, which should be carried out in each phase.

The main outcome of the initiation phase is the feasibility study, which should answer following questions: (i) Is KMS needed at all? (ii) What kind of KMS is best suitable? The requirements specification phase involves preparing a detailed requirements specification for all three categories. For specifying application and access interface requirements common software engineering techniques should be used. For specifying ontology requirements, a ontology requirements specification document should be created (ORSD). The structure of the ORSD is described in detail by the NeOn methodology.

Another important technique for developing ontologies are competency questions (CQ). They specify questions to which the final ontology (hence knowledge base) should be able to provide answers and are defined in ORSD. CQ are formalized in the modelling phase and are also used for ontology testing and vali-
5 Conclusion

In this paper we have analyzed approaches for deployment of KMS in organizations. There is a significant difference between common and intelligent KMS. Intelligent KMS provide much better knowledge management functionality, however knowledge has to be represented in a machine readable manner. Because of this, a lot more effort is needed to develop and deploy intelligent KMS. We conclude that in case advanced knowledge management is needed (e.g. machine readable knowledge, machine processing of information, reasoning), then the most suitable approach is ontology based KMS, otherwise appropriate common KMS should be chosen. In the second part of the paper, we have proposed an architecture and a method for developing ontology based knowledge management systems, which are on one hand capable of machine readable knowledge representation and on the other hand interoperable with legacy applications and modern programming languages.

References:


