Journal: Journal of Information & Knowledge Management,
Vol. 3, No.1 (March 2004), pp. 47-66

Title: **h-TechSight – A Next Generation Knowledge Management Platform**

Authors: Michael Stollberg, Anna V. Zhdanova, Dieter Fensel

Affiliation: DERI – Digital Enterprise Research Institute
University of Innsbruck, Austria, and
National University of Ireland at Galway, Ireland

{michael.stollberg, anna.zhdanova, dieter.fensel}@deri.ie

Corresponding Author: Michael Stollberg

Address: Leopold-Franzens Universität Innsbruck
Institute for Computer Science
Technikerstrasse 13
A-6020 Innsbruck
Austria

E-mail: michael.stollberg@deri.ie
Tel.: +43 (0) 512-507-6479

Keywords:

Knowledge Management, Next Generation Knowledge Management, Ontologies, Semantic Web, Natural Language Processing, Knowledge Discovery

TABLE OF CONTENTS

h-TechSight – A Next Generation Knowledge Management Platform

Michael Stollberg, Anna V. Zhdanova, Dieter Fensel

DERI – Digital Enterprise Research Institute, University of Innsbruck, Austria

ABSTRACT

A new breed of Knowledge Management Systems (KMS), referred to as the Next Generation of Knowledge Management (NGKM), aims at overcoming the limitations of existing KMS regarding the range and quality of supportable KM activities. Therefore, AI-based technologies are employed to enable automated knowledge item processing instead of the information retrieval functionality provided by conventional KMS. This paper presents the Knowledge Management Platform KMP developed in the h-TechSight project as an NGKM-solution. It is designed to support organizations in technology intensive industries in monitoring changes in their external environment as an important factor in competitiveness. Therefore, the KMP combines NGKM-enabling technologies for automatically observation of information resources on the Internet in order to notify users whenever a change in their domain of interest occurs. The aim of this paper is to position the KMP as a KMS of the next generation and to delineate the lessons learned in KMS design in the h-TechSight project. Thus, we illustrate the idea of NGKM with special emphasis on enabling technologies and, on this basis, expose the functional design and technical realization of the h-TechSight KMP.

Keywords: Next Generation Knowledge Management, Knowledge Management Platform, Ontologies, Semantic Web, Natural Language Processing, Knowledge Discovery

1. INTRODUCTION

Since knowledge has been discovered to be a competitive asset of major importance, a large number of general approaches and supportive technological systems for Knowledge Management (KM) have been developed. In general, organizational KM aims at providing support for managing information and knowledge management inside organizations and between them which is determined by several aspects. A profitable KM-solution for a specific organization has to capture, integrate and maintain information important for the organisation, and it has to provide knowledge transfer and application support in a way that corresponds to the organizational culture. Although IT-technology is not mandatory for creating a valuable KM-solution, it can increase the solution's functionality tremendously. The requirements for such technology suites, called Knowledge Management Systems (KMS), are very complex and multifaceted. One the one hand, a KMS has to be customizable for the KM-requirements of a specific organization; on the other hand it has to provide technical support for handling information in a sufficient manner.

Existing KMS offer different functionalities for KM-support, which can be classified into systems for data management and retrieval, for communication and messaging, and for information browsing and retrieval (Alavi and Leidner, 1999). A substantial drawback common to these KMS is that they use conventional IT-technologies which are limited in their abilities to process information. More precisely, this means that the technologies that constitute the basis of KMS allow management of documents and other information but do not support automated information processing like, for instance, automated detection of relevant information for a certain topic or extracting new information out of electronic resources. Thus, maintenance of the information held in KMS is still left to human intervention, and, resulting from the primary strength of conventional IT-technologies, many KMS merely offer information retrieval functionalities like indexing, organizing and retrieving documents and information. A new generation of KMS, frequently referred to as the Next Generation of Knowledge Management (NGKM), aims to overcome these deficiencies by using AI-based technologies as the technical foundation of KMS. These technologies improve the functional power of KMS in two ways. At first, they allow automated maintenance of information in a KMS instead of human intervention. Secondly they support automated extraction, analysis and processing of knowledge, thus extending the information retrieval-oriented functionality of a KMS by automated knowledge-item-processing facilities which offer a more convenient support for KM-activities (Maedche et al., 2003).

The objective of the h-TechSight project[1] is to develop the Knowledge Management Platform KMP along with a KM-methodology to improve capabilities of technology intensive organisations to monitor, assess, predict and respond to trends and changes in their domain of activity. The scientific aspects researched in the h-TechSight project are, on the one hand, KMS design requirements for the specified field of application, and, on the other hand, the technological development of the KMP using NGKM-technologies such as ontologies, agent technology, and intelligent web search. The outcome of the project will be technological know-how for implementation of NGKM-technologies as well as insights on the design and applicability of next generation KMS.

The goal of this paper is to indicate the requirements for next generation KMS and present the combination of NGKM-technologies implemented in the h-TechSight KMP, and to also point out practical issues on KMS design disclosed in the h-TechSight project. In order to allow a positioning of the KMP as a NGKM-solution we first address the objectives of KM and discuss the challenges arising for KMS of the next generation with special attention to the enabling technologies, then expose the design and technical realization of the h-TechSight KMP and point out future obligations for the development of NGKM-solutions.

The paper is organized as follows: section 2 reassembles the general goals and major concerns of KM and presents the requirements arsing for NGKM. Section 3 examines the technological demands and exposes three general groups of NGKM-enabling technologies. Section 4 depicts the specific knowledge management approach followed in the h-TechSight project, inspects the functional design as well as the technical realization of the KMP and summaries the research results achieved. Related work is examined in Section 5 and finally section 6 concludes the paper.

1 h-TechSight Project (IST-2001-33174): "A knowledge management platform with intelligence and insight capabilities for technology intensive industries". See website at: http://prise-serv.cpe.surrey.ac.uk/techsight

2. EMERGING CHALLENGES IN KNOWLEDGE MANAGEMENT

This section reassembles the challenges arising in KM and KMS development currently discussed in research and development activities in order to attain a theoretical basis for further investigations. We replicate the notions, concerns and objectives underlying organizational KM and point out the functional requirements arsing for modern KMS, which naturally also hold for NGKM-systems.

2.1 Foundations of Knowledge Management

As a steady increase in the percentage of people employed in information-related fields has been reported during the last several decades, the interest in KM is rising and an efficient treating of this tendency is considered to be important for an organization's competitive ability. In order to provide a profound foundation for further investigations, we need to clarify the concepts underlying KM and KMS. As this has been treated extensively in literature (Alavi and Leidner, 2001; Bhatt, 2001), here we only briefly recapitulate the understandings of knowledge, KM and KMS.

For understanding the objectives and difficulties of KM, we first have to elucidate *what* is to manage, i.e. understand what knowledge is. Although an ultimate understanding of knowledge does not exist and different, sometimes oppositional views on knowledge in literature of various scientific disciplines can be reported, there is a commonly accepted delineation of what knowledge is in the terms of KM. Therein, knowledge describes what enables a human to carry out a task or to make a decision. More precisely, knowledge is information which is applied to solve a certain problem; it gives skills or competence to a human. A fundamental property of knowledge in this context is that it is always adherent to an individual human, further it has to be learned by each individual and can not simply be transferred into another human being. This means that for transferring knowledge, it has to be transformed into a distinctive code which can be communicated, then be decoded and learned again by the receiver in order to gain the same competence.

An important aspect of understanding of knowledge in KM is the differentiation of explicit and implicit or tacit knowledge (Polanyi, 1967). The former describes schematic, deterministic knowledge which can be articulated and thus be transferred very easily. The latter describes knowledge which is intangible, which means that only its owner can use it to solve problems but it is very complicated to explicate and transfer this kind of knowledge. Nonaka and Takeuchi detected the importance of tacit knowledge for organizational KM and introduced their widely known "Theory of Knowledge Creation" (Nonaka and Takeuchi, 1995), and numerous research activities deal with this problem, e.g. (Stenmark, 2001). Another important conceptual aspect on knowledge in KM is the differentiation of data, information, and knowledge relying on a semantic model of communication. Therein *data* is understood as un-interpreted signals without application, i.e. the syntactic level; *information* is data in an application scenario, so it is already equipped with meaning on the semantic level; and *knowledge* is information used for solving specific problems, thus representing the pragmatic level. This model derives from a technical point of view, and is mostly used in Knowledge Engineering systems to describe the processing features of a Knowledge System (Schreiber et al., 2000).

These two aspects can be seen as the two dimensions of knowledge which have to be dealt with in KM. Figure 1 illustrate this.

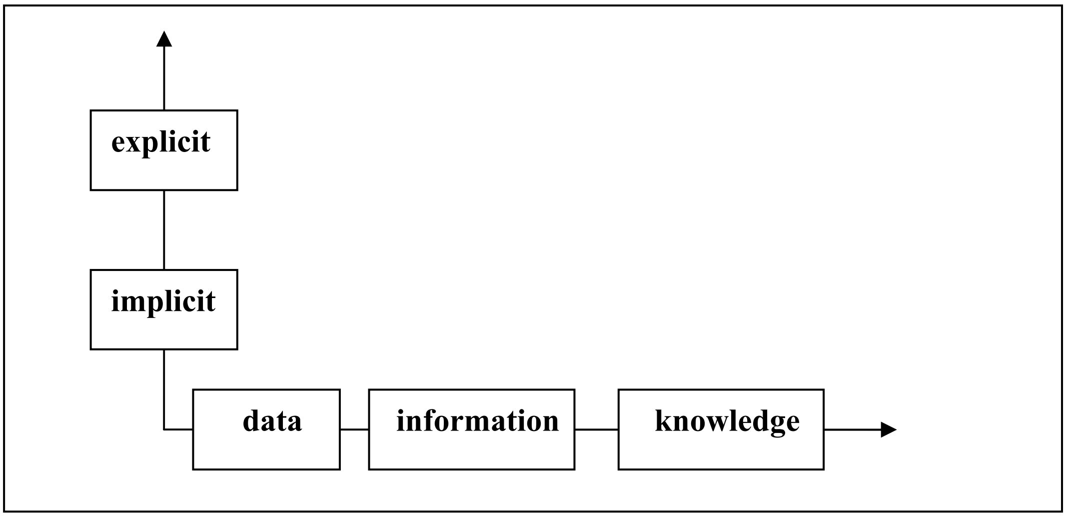

Figure 1: Understanding of "knowledge" in KM

The objective of KM is to make knowledge, the explicit and the tacit knowledge existing inside an organization as well as the relevant external knowledge, usable in an organization and obtain competitive benefit from this. Broadly speaking, the aim of KM-solutions is to provide tools, techniques and processes to enable the most effective and efficient utilization of knowledge in an organization. The main challenges to be treated in KM is to turn data into information (i.e., placing it into a sensible application scenario), to make information available for organization members via an appropriate dissemination technique, and to support people to gain knowledge via learning. The construction of a sophisticated KM-solution for a specific organization heavily depends on the information needs as well as on the organizational culture of the organization. Therefore, KM-methodologies have been developed which aim at defining a general framework for setting up KM-solutions. Therein typically a set of interrelated actions is defined. Table 1 summarizes the subsequent KM-activities which represent the very general steps (Davenport and Prusak, 1998).

Table 1: Knowledge Management Process

KM-Activity	Description
1) **Knowledge Identification**	understanding the character of the needed knowledge, picking out existing relevant knowledge, and allocating the knowledge assets which need to be learned and created
2) **Knowledge Acquisition**	obtaining the needed knowledge (e.g., buying / consulting, R&D, learning / self-creation)
3) **Knowledge Preparation**	presentation of information in an easy to learn way (e.g., explicit knowledge can be turned into documents, and implicit knowledge can be reflected into contact information / human expertise)
4) **Knowledge Allocation**	presentation of information in an easy accessible way

5) Knowledge Dissemination	creation of a knowledge distribution infrastructure
6) Knowledge Usage	ensuring that people use knowledge in KM system
7) Knowledge Maintenance	keeping a KM system in an up-to-date condition

Finally, a Knowledge Management System (KMS) is an IT-infrastructure designed for supporting the KM-process by tools and technologies for electronic management of knowledge or information, respectively. The design of sophisticated KMS is a very complex and challenging task as a KMS has to adapt to the specific KM-requirements of an organisation. This means that it has to support the concrete tasks of creating, allocating, disseminating, and maintaining knowledge and it should add surplus value for the KM-activities in an organization. Design guidelines and the associated problems have been widely studied in literature, concluding in frameworks for classifying different types of KMS and identifying success factors for KMS (Hahn and Subramani, 2000).

2.2 Functional Requirements for Next Generation Knowledge Management Systems

As mentioned introductory, current KMS lack of support for several aspects important for KM, both functional deficiencies as well as limitations of the underlying technologies. Before we will discuss the technology reorientation in the next section, we outline the functional requirements for next generation KMS in the following.

The Internet has gained importance as a large, world-wide information space and information exchange medium. As information relevant for organizational KM can be found, modern KMS approaches are supposed to make use of the Internet as an information repository (Applehans, 1998). From the functional point of view, recent KMS have been detected to be insufficient with regard to the support offered for KM-activities, thus not adding profitable value for KM. McElroy discussed these shortcomings extensively and identified eight functional requirements for the next generation of KMS (McElroy, 2002). We will briefly summarize these aspects because they can serve as a classification framework for functionality of KMS.

Supply and Demand Sides of KM

Currently, a major part of the KM-technologies such as groupware, information indexing and retrieval systems, knowledge repositories, document management, etc. realize a ´information-pull´-approach for knowledge identification and transfer. Such strategies which are based on the distribution of already *existing* knowledge represent the supply side of KM. The demand side of KM is concerned more on accelerating the production of *new* knowledge requested by users (i.e. ´information-pull´). Supply side KM is relatively well-developed while there are only very few solutions for demand side KM. Thus, support for this (i.e., high-performance learning throughout the whole KM-process) is a major challenge for NGKM-solutions. It is appropriate to note that demand-side strategies deal with the whole knowledge work cycle and not just its downstream events.

The Knowledge Life Cycle

Current KMS are concentrated on codification and transfer issues, omitting the fact that organizations do not only hold collective knowledge but they constantly evolve and learn. NGKM-solutions should support the whole knowledge life cycle, i.e. knowledge *production*, *validation*, and *integration* in organizations.

Knowledge Processes

A new aspect is that NGKM should also be concerned with knowledge process management. The perspective of work with business processes offers the possibility to identify knowledge sources and products as well as best practices and knowledge deficits in the business processes and, hence, optimize these processes. Although there are applications which connect information retrieval tools and data warehousing with the knowledge worker requirements, the linkage to the working processes is still unsatisfactory.

Knowledge as Rules

It is the knowledge of an organisation that defines its practical behaviour. We do what we do because of what we believe (our *know-what* or declarative knowledge), and we do what we do the way we do it because of our *know-how*, also called procedural knowledge. These kinds of knowledge can be expressed in *rule sets*, wherein knowledge induces action. A challenge of NGKM is to find ways of capturing and expressing organisational rules.

Knowledge Structures

Knowledge in an organization can be described by particular knowledge structure, which covers a specific kind of knowledge like organizational structure, business strategies, products and services, business processes, etc. An essential shortcoming of current KM-technologies is the lack of appropriate techniques for representing and managing knowledge structures. In NGKM-systems, the different knowledge structures should be conceptually presentable and technologies for processing and transferring should be provided.

Nested Knowledge Domains

The concept of nested knowledge domains points to differentiated levels of knowledge relevant for organization KM: knowledge held by individuals, knowledge held by groups of individuals, knowledge held by the organization as a whole, and knowledge held in a domain of interest between organizations. At any point in time disparities are bound to exist between these levels, thereby prompting disagreements and different points of view. However, out of this creative tension between the knowledge domains new knowledge is derivable. The challenge for NGKM-solutions is to manage this tension effectively in order to boost the rates of innovation and organizational performance.

Organizational Learning

As noted above, organizations do not only have to manage the knowledge they already posses but also they have to adapt to changes in their external environment in order to keep up competitiveness. Nowadays, the explicit connection is being drawn between KM and organizational learning (OL). Thus, NGKM-solutions should be designed and recognized as an implementation strategy for OL.

Complexity Theory

Nowadays the science of complexity is seen as a valuable sort of insight in the understanding of how living systems behave. In particular, the Complex Adaptive Systems (CAS) theory holds that living systems (i.e., organisations made up of living, independent agents such as people) self-organize and continuously fit themselves to the ever-changing conditions of their environment by modifying their knowledge in fact and in practice. Thus the CAS theory can be considered as the direct source of thinking behind the notions discussed above (Wilensky and Resnick, 1999). It is an important requirement for NGKM to provide means for dealing with the environmental complexity present in organizations.

Concluding this subsection, it is important to point out that KMS can not be realized by only building a technological KM-infrastructure because knowledge is more than just data or information as it is always attributed to a human individual. A large amount of non-technical effort is needed to create a successful KM system in an organization. However, technologies enable KM as they potentially can provide support for most of the functional requirements listed above, and further the technology employed determines the quality of the whole KM system. The more of the NGKM-requirements technologies support in an adequate way, the better a KMS built upon them can appreciate the KM objectives. Therefore, the role of technologies remains to be of primary importance for the KMS of the future.

3. Next Generation Knowledge Management Technologies

The technologies employed in KMS essentially determine the application quality of the system. As stated introductory, conventional IT-technologies lack of capabilities for automated knowledge-item processing which promises much more convenient support for KM than information retrieval facilities.

The idea of NGKM is to replace conventional IT-technologies by AI-based technologies as the technological basis of KMS (Smith and Farquhar, 2001). Especially in AI-research fields like Knowledge Representation and Knowledge Engineering are concerned with technologies to codify and automatically process knowledge items (Studer et al., 1998). The problems addressed in this fields are comparable to the requirements pointed out for KMS, i.e. how to represent and structure knowledge and how to identify the relevant knowledge needed to solve a certain problem. The aim of research in these fields is to develop technologies for automation of these processes in order to move activities like searching and interpreting information to the computer, thus decreasing the need of human intervention for knowledge detection and maintenance. As such functionalities are also considered in KM-approaches, the use of AI-based technologies as the basis for KMS promises more expedient support for KM-activities (Milton et al., 1999). Moreover, AI-technologies enable a better support for the NGKM-requirements for KMS discussed above.

Three groups of technologies that derive from AI-research have been identified for enabling NGKM-solutions – especially in the context of web-enabled KM (Abecker et al., 2003). (1) Enhanced Knowledge Representation technologies to provide more expressive knowledge modelling and representation technique, (2) Natural Language technologies for automated text processing, and (3) Knowledge Discovery technologies for automated information processing which allow to implement “business intelligence” (Kalakota and Robinson, 2000) relying upon (1) and (2). These three groups describe the foundation of NGKM-systems from a technological point of view. Currently, these technologies have only been identified as potential building blocks for NGKM-solutions but no sophisticated realization in a real-world KMS is to be recorded at this point of time (Anagnostou et al., 2003). In the following we discuss these technologies, focussing on the resulting benefits for KM.

3.1 Knowledge Representation

Knowledge Representation (KR) comprises techniques for modelling and representation of knowledge structures in a way that is readable for humans and as well processable by machines. Further, a KR technology allows a high extent of expressiveness in order to represent knowledge structures in an appropriate way. As stated above, an essential drawback of first generation KM-solutions was the absence of sufficient KR mechanisms. Within current research efforts ontologies are favoured as a satisfactory KR technique. The reason for this is that ontologies provide, on the one hand, a means for human-readable knowledge modelling not determined by technological enforcement and on the other hand machine-processable semantics. Here we briefly introduce ontologies and show their benefits for KM, especially for ontology-enhanced web search (Russell and Norvig, 2003, pp. 320 ff).

Ontologies

In Artificial Intelligence an ontology is understood as a “formal, explicit specification of a shared conceptualization” (Gruber, 1993). This means that an ontology can be seen as a conceptual model of a domain which describes the semantic – i.e., describing a sensible context – knowledge structure of the domain. Furthermore, the creation of an ontology is aimed at reaching a commonly accepted consensus in understanding of the domain among the multiple participants involved. The descriptions of the knowledge structure should be explicit, i.e., every aspect should be modelled precisely in the ontology without skipping information which might be considered as commonly known. Thus, a terminological definition of the domain should be achieved to improve the understanding of a domain and simplify communication. Several methodologies have been developed to facilitate the creation of ontologies (Fernández, 1999).

Furthermore, an ontology can be specified in a formal language which enables a machine to process the ontology. To enable exchange of ontology data over the Internet, web-compatible ontology specification languages have been developed which are referred to as Semantic Web languages (Gómez-Pérez and Cocho, 2002). An ontology thus provides formalized semantics and can be applied as the underlying data model of an information processing application. Using semantically enhanced descriptions of information can significantly improve the quality of information processing of a system. As ontology technologies are considered a key component for NGKM solutions, we will summarize the general benefits that ontologies provide to the KM-technology.

First, an ontology provides an unambiguous terminology definition. Thus it allows overcoming misunderstandings in a knowledge domain and communication problems caused by controversial vocabulary used in a domain. Furthermore, ontologies can be referred to as a shared understanding of a knowledge domain which people can refer to as a common conceptual model which facilitates identification of relevant information as well as coherent information exchange in a community. Secondly, ontology techniques provide richer expressiveness than conventional modelling techniques (like ER-Diagrams or Object-Orientated techniques, for example) for describing knowledge structures and they provide a conceptual model instead of a technical motivated depiction of a knowledge domain. Thirdly, an ontology provides formally specified, machine-processable semantics which can be utilized as the underlying data model of a knowledge processing application. Thereby the quality of information processing can be increased significantly. In addition, ontology maintenance technologies enable evolvement of an ontology during system runtime, thereby supporting assimilation of changes in a dynamically changing domain of knowledge. Last but not least, ontologies can be employed broadly in information structuring through aligning information items as instance data to the ontology (Abdecker and van Elst, 2003).

Regarding the NGKM-trends presented above, ontologies provide means for defining knowledge structures and processes, as well as nested knowledge domains by representing them as networks of ontologies. Also, ontologies facilitate organizational learning as the knowledge of an organization is brought to attention and evolving ontologies allow tracking the changes in this knowledge. Finally, ontologies enable reduction of complexity since they allow structuring and explicitly describing information structures.

Ontology-enhanced Web Search

The increasing importance of the Internet as a valuable information repository for KM has been outlined above. The most important information processing feature with regard to the usability of the Internet as an information resource for KM-solution is information retrieval, i.e. web-search technologies. Therein the general problem is how to retrieve the information that a search request is really looking for – i.e. if you enter the search term “Jaguar” you will receive information on the animal as well as on the automobile as a search result. Current web search engines like AltaVista, Goggle, etc., use very complex search mechanisms to increase the quality of their web search results but still retrieve a lot of irrelevant information.

Ontologies and Semantic Web technologies aim at overcoming these weaknesses of the current web. In the Semantic Web, information on the web will be semantically annotated via ontologies as a predefined knowledge structures. Sticking to the example considered above, a search query for the term “Jaguar” will be semantically classified. If the term “Jaguar” is depicted as a sub-concept of “Animal”, the web search results can be restricted to web content that is related to the animal jaguar via the properties defined for this ontological concept. Thereby the precision ratio (number of relevant retrievals in relation to the total number of retrievals) of web search techniques can be increased tremendously. This will allow a much more sophisticated usability of the Internet as an information repository (Fensel, 2003).

Currently, only a very small amount of web content has been semantically annotated so that the Semantic Web enhanced web search functionality can not be employed throughout the whole Internet yet. But ontologies can also be used to increase the quality of web-search on “ordinary”, not semantically annotated web content. If the search term “Jaguar” is related to the sub-concept of animal, the properties defined for the ontology concept can be used as additional search information. Thereby, the precision rate can also be increased, although not to that extent which is achievable with ontology-based search on Semantic Web content (Krohn and Davis, 2001).

The reasons why ontology technology is a key-enabling technology for NGKM within the context of web search can be summarized as follows:

- Ontologies provide a knowledge modelling technique with high expressiveness which allows to represent the structure of a knowledge domain explicitly

- As ontology development is a collaborative process of the community members involved in the domain, an ontology can be referred to as a commonly accepted understanding of a domain which helps overcoming communication problems and facilitates reduction of complexity and thus enables organizational learning
- A formally specified ontology can be employed as a grounding data model of information processing applications which provides formal semantics of information, and thereby can enhance the information processing quality of a system
- Ontologies and Semantic Web technologies allow overcoming the shortcomings of current web search technologies and thus strengthen the usability of the Internet as a worldwide information repository which is an important factor of NGKM-solutions.

3.2 Natural Language Processing

A considerable amount of data on the Internet is and will be represented in natural language. Further, most likely people will go on using natural language as a primary mean for information exchange. Therefore language technologies which help to process (web)data to information and information into knowledge, and provide a faster and more effective human-computer-human interaction naturally lie in the core of the NGKM-technologies. The discipline for these language technologies is Natural Language Processing (NLP) which studies analysis of natural written or spoken language plus its application thereof. The applications include such pragmatic aspects as natural language interfaces, machine translation, natural language understanding and generation, dialogue and discourse systems, e-mail classification (e.g., spam filtering), computer-aided language learning and text summarization.

The NLP activities within the language analysis take place at 7 levels (Allen, 1994) and (Engels and Bremdal, 2000). To demonstrate the contribution of NLP-technologies for enabling NGKM, we will list these levels and the processes corresponding to them in the following.

Phonological level is concerned with speech recognition and generation. The activities at this level do not occur often in KM, since usually only little content is presented in audio files or streams. However, NLP technologies at phonological level can be employed for providing human-computer interaction at the places where visual media is inconvenient or impossible to use (e.g., surgery).

Lexical level. The processes of the lexical level take natural language texts as an input. The text can pass a tokenization phase where lexical units (paragraphs, sentences, words, etc.) are identified and be labelled with part-of-speech tags. In some statistical approaches of NLP (e.g., n-gram approaches for e-mail classification and information retrieval), the tokenization phase can be very simple and the tagging process can be omitted. The activities on the lexical level are essential for all word-based approaches in NLP.

Morphological level includes resolution of suffixes, prefixes and inflectional forms of words, compound analysis and, finally, stemming. Stemming is a process that transfers a word-form into its primary form or root. For example, the verbs *carried*, *carries*, *carrying* can be transferred to their primary form *carry*. With compound analysis one can split compound words into their basic components. E.g., a German word

Hauptbahnhof can be transferred to the words *haupt*, *Bahn* and *Hof*. Activities at morphological level are important for web search and linking ontologies with instance data, because they allow establishing co-references between the same words written in different ways.

Syntactic level incorporates two major tasks. The first one is sentence-fragment identification and the second one is assigning the exact roles to individual words, taking into consideration the grammatical rules of a language and a description (grammar) of how words can be put together in that language. The activities performed at this level are helpful for ontology schema and instance data retrieval and storage. The contribution of these activities is the new relationships between concepts, classes, instances and attributes which can be derived by using grammars and statistical parsing techniques.

Semantic level is covered by three major topics: knowledge representation, word sense disambiguation and extension of knowledge representations with synonyms or related words. The process of adding semantics to a representation of some natural language text requires a sufficient knowledge representation formalism to be defined. An ontology is an example of an explicit knowledge representation formalism with formally specified, machine-processable semantics; therefore ontology technology fits tightly the semantic level of NLP. Thus, in our context, word sense disambiguation is to be seen as choosing correct co-references between the ontology elements among the multiple possible ones. The NLP-techniques for identifying synonym groups are also directly related to knowledge discovery by ensuring that the concepts, which convey the same meaning but are expressed in different linguistic ways, are aligned similarly to an ontology.

Discourse level. Understanding of language significantly depends on a context. While the semantic level introduces a first semantic analysis (mainly on the sentence level), the level of discourse takes into consideration the context on a larger scale, mainly based on the experiences and evidence from the whole narrative. For example, consider a sentence taken out of some text: “Now he just realized what was going on”. One can see that making pronoun resolution and ascribing this sentence to some of the known contexts are necessary for turning the data of the sentence into information. In a case like ours, the discourse level activities basically assure that information is properly interpreted in ontology filling.

Pragmatic level of NLP incorporates resolving all utterances, ambiguities and utilizing as much world knowledge as necessary (or available) for doing so. Here, the NLP-techniques are aimed at interpreting intentions, intuition, and metaphoric expressions employing common sense knowledge. Actually, the NLP processes at the pragmatic level lie in the base of turning information into knowledge. According to the NGKM concept of nested knowledge domains, the semantic representation of a domain can vary depending on whether it belongs to an individual, group of individuals or the organization as a whole. NLP is capable of providing the means for treating a semantic representation at any levels.

In a nutshell, the importance of NLP-based technologies is based on the fact that they release the reader of the burden of extracting and interpreting the relevant information found in natural language texts. Thus, the NLP-technologies are the grounding technologies for KM, because of their substantial contribution to:

- automated web content processing
- machine-processable annotation of natural language texts with ontological notions
- discovering new ontology elements (i.e., concepts, classes, instances, attributes, relations, claims) for the purpose of aligning them to ontologies
- increasing recall and precision of the web search engines.

Thus, NLP is indispensable for the advanced and large-scale automated web content processing applications of the NGKM.

3.3 Knowledge Discovery

Upon ontology technologies which provide the underlying, machine-processable semantics of data and NLP-technologies, which support automated text processing, the third group of enabling technologies for the NGKM solutions is Knowledge Discovery technologies (KD). KD is a research area of Artificial Intelligence applied to extract potentially interesting and previously unknown knowledge from large quantities of data and information, providing technologies for semi-automatic knowledge extraction and acquisition.

The disciplines related to KD include data-, text-, and web mining, machine learning, and information extraction/retrieval. These techniques employ statistical approaches, symbolic and machine learning and classification techniques like document clustering, conceptual graphs, decision trees, neural networks, self organizing maps, genetic algorithms, association rules and case based reasoning (Fayyad et al., 1996). Possible employment scenarios of KD technologies range from search engines to web portals and personalization modules. For example, preference-based personalization approaches typically rely on some notion of web mining: web content mining for exploiting the content of hypertext documents, web structure mining for analyzing the hypertext structure, or server / browser logs as web usage mining techniques. In a nutshell, KD technologies provide automation for extracting relevant aspects out of information resources and analyzing information according to pre-defined preferences, thus enable automated knowledge identification and acquisition.

The applicability of KD technologies in NGKM-systems based on ontologies and NLP-techniques is interrelated with the knowledge structures defined in the ontology as well as on the outcomes of the NLP-tools. A KD-process has to rely on a common knowledge structure, i.e. the ontology, supplemented with relevant instances, regularly derived by NLP-techniques. As KD facilities are built upon the other components of a NGKM-system, their functional design has to be adjusted to the ontology structure as well as to the NLP-features in order to provide sophisticated functionality. Thus, the design of the technical components underlying the next generation of KMS is as complex and multifaceted as for conventional KMS (Anagnostou et al., 2003). Nevertheless, KD-technologies are a means for automated knowledge-item processing and thus are the essential component to exploit the full potential of NGKM-systems. Especially in the context of web-based KM, KD technologies enable semi-automated web content processing and thus enable the use of the Internet as a world wide information repository not only for human consumption but also for automated information processing systems (Kosala and Blockeel, 2000).

Closing this section, we summarize the facilities that KD-technologies provide for NGKM-systems:

- semi-automated web content processing, which helps to avoid information overload

- simple access to data available from a multiple of diverse autonomous distributed resources, which contributes to effective knowledge sharing, better decision making, development of new knowledge areas, and saves the time spent on searching for relevant knowledge resources
- support for information and knowledge item organization (e.g., aligning web content to ontologies as instance data), which allows to derive tacit knowledge, capture it and make it explicit
- automated analysis of business processes and business practices, which helps to improve these processes as the dynamic part of KM
- capturing of the existing domains and inter-domain relationships and exploring possibly new relationships, which accelerates innovation processes and organizational learning by detecting new knowledge areas as the static part of KM.

4. h-TechSight Knowledge Management Platform

In the previous sections we have outlined the emerging challenges in KM and requirements for next generation KMS. On this basis we will present the design approach and technical realization of the Knowledge Management Platform KMP developed in the h-TechSight project. We explain the approach followed for designing the KMP and then describe its technical realization in order to indicate the contributions of the project outcome in terms of findings on KMS design and on technical requirements for future NGKM-solutions.

4.1 Project Objectives and Approach

The h-TechSight project is a 5th framework project of the Information Society programme funded by the European Commission. The project consortium consist of 11 partners, assorted into technology providers, comprised of commercial companies and research institutes, and a users group which consists of SMEs and large-scale enterprises working in the chemical industry. The development of the KMP is based on a requirements analysis of the KM-needs for technology intensive industries, technological objectives for realizing a NGKM-solution, and an iterative test-refinement process within the project consortium.

Objectives

The overall objective of the h-TechSight project is to provide a complete KM solution to enable companies in technological intensive industries, especially SMEs, to monitor, assess, predict and respond to knowledge-related changes in their external environment in a better way. In organizational management, the external environment is understood as the context an organization acts in, comprised of the industrial sector, its competitors, and socio-economical regulations. This has been identified to be of major importance for enterprises with regard to competitiveness and market positioning, thus influencing strategic directions to a very high extent (Nickell, 1995). In terms of KM, up-to-date information on the external environment of an organization is considered to be very important as it essentially determines strategic decisions – especially in technology intensive industries, which by nature are also very information intensive. As the Internet has been identified as an important information source for modern KMS, and possibly be of even higher importance for technology related industries, the h-TechSight project aims at creating a KM-solution for automatic monitoring changes in internet resources which contain relevant information on the external environment of organizations working in technology intensive industries (Fernando and Kokkosis, 2003).

This objective is displayed in the functional design of the KMP. It allows users, i.e. an organization, to characterize its external environment by defining an ontology or by selecting a pre-defined ontology that describes the domain of interest. Then the system automatically monitors internet resources that hold relevant information by means of ontology-enhanced web search and web mining. If a change occurs in these internet resources, this will be reported and pre-processed in the h-TechSight KMP, and a user who has selected this domain to be monitored as a valuable knowledge resource for his business will be notified immediately. Technically, the KMP employs NGKM-technologies as described above. Ontologies are used to define the knowledge structure of the knowledge domains covered, an agent-based search engine is developed for intelligent information retrieval from the web as well as ontology-enhanced web mining, and knowledge discovery technologies are used for analyzing and pre-processing information retrieved from the Internet. The technical realization will be investigated in the following section in more detail.

The outcome of the h-TechSight project will be, besides the KMP, sample ontologies for identified knowledge domains which will serve as a starting point and be refined during the system's runtime, assimilation of user requirements, especially from the user group – partners in the project consortium, and a fine-grained acquaintance on KM user requirements as well as methodological know-how for implementing NGKM-solutions. We will discuss the results of the project in detail in section 4.3.

Approach

As already mentioned, the development of the KMP is a three-phase process. First, the needs for KM in technological intensive industries have been detected via a representative survey on knowledge management activities and requirements in this industrial section. The most important results on why KM technologies are considered to be essential and required functionalities have been (1) environmental pressures, i.e. faster and more competitive markets, (2) the web gains more importance as an information resource, and (3) automated technical support for creating valuable knowledge is considered to be important. A further important outcome has been that currently used KM-solutions, especially in SMEs, are proprietary, manual systems. The reason for not employing existing KMS is that these mostly are monolithic and costly systems which, one the one hand, would overwhelm the KM-support needed and, on the other hand, exceed the available KM-budget (Fernando et al., 2002). On basis of this as well as on exhaustive studies on the state of the art in NGKM-technologies, the technical design and implementation of the KMP is done in the second phase of the project. In the last phase, the user group of the consortium performs an extensive usability evaluation based on a pre-defined, detailed evaluation procedure. In a subsequent iteration step, the KMP will be refined according to the test results and usability evaluation.

This development process includes all groups which are important for creating a profitable KMS: the KM-requirements arising in user community, the technology developers and a representative group of end users for usability evaluation of the system. The close collaboration in the project consortium and the iterative refinement during development ensures that the KMP fulfils the functional requirements targeted at. Thus, with respect to different development approaches identified in KMS development, the h-TechSight approach follows an action-case – methodology (Braa and Vidgen, 1999).

The special focus in the following sections is on the technical realization of the h-TechSight KMP with regard to the NGKM-enabling technologies as depicted in section 3.

4.2 KMP Architecture

The Knowledge Management Platform KMP consists of three major components. First, the Dynamic Ontology Management System (short: DOMS) which stores and retrieves the ontology schema and instance data of the KMP. Second the Search Module which consists of a Multi-Agent Search Engine (short: MAS) which performs a general ontology-based web search and a Web Mining tool for monitoring specific web resources. The third component is the Knowledge Discovery Module which analyses the search results retrieved of the Search Module, pre-processes this information and aligns them as instance data to the corresponding ontology. Figure 2 presents an overview of the KMP architecture. The workflow and the interoperation of the components will be explained in more detail below.

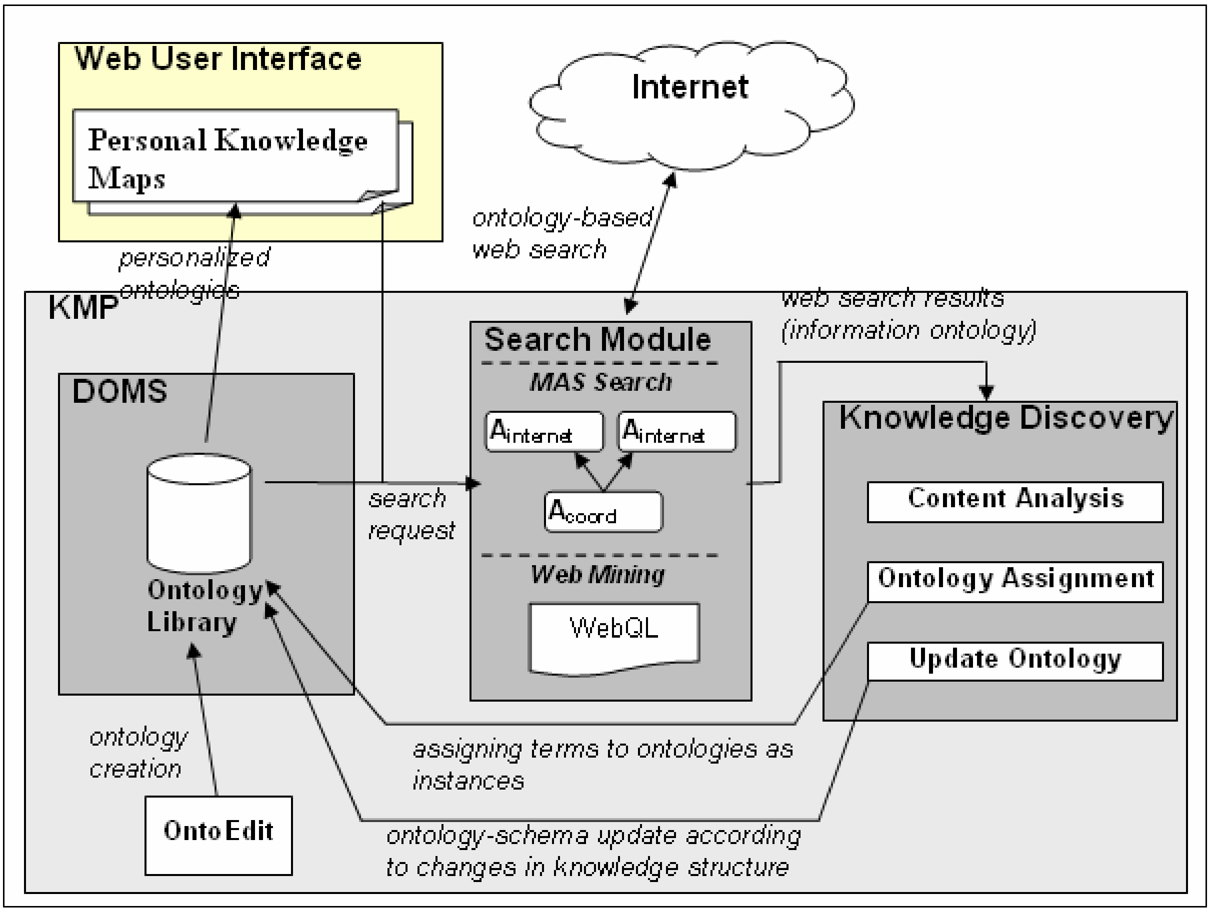

Figure 2: KMP Architecture

The KMP is based on the ontologies stored in the DOMS, the ontology library of the system. The initial ontologies have been developed in collaboration of the h-TechSight consortium and domain experts outside the platform using the OntoEdit[2] ontology

[2] OntoEdit: ontology editor developed by Ontoprise GmbH (see: www.ontoprise.de, introduction in (Sure et al., 2002a). For ontology creation in h-TechSight, the professional version of OntoEdit has been used.

editor, whereby RDF [3] is used as the ontology representation format. The User accesses KMP via a web interface, thus allowing remote usage. As not the complete ontology might be of interest for every user, each user can specify a personal knowledge map which is represented by a sub-tree of the ontology hierarchy. For this personalized ontology a user can initialize a web search. Therefore the requested ontology sub-tree is passed to the Search Module. Therein, depending on the search mode chosen, either the Multi-Agent Search Engine MAS performs a web search for the selected terms or the Web Mining tool inspects specific web resources for the search request. The search result consists of a so called ´information ontology´ enclosing web resources that contain relevant information for the ontology sub-tree searched for. This is passed on to the Knowledge Discovery Module for further processing. Therein, at first the content of the web resources is analyzed in order to retrieve the information searched for, secondly the detected terms are assigned to the corresponding ontology and stored in the DOMS as instance data, and thirdly, potential changes in the knowledge structure are detected in respect of the existing ontology. These changes will be transmitted into the ontology as an ontology schema update. Thus, the h-TechSight KMP implements a dynamically evolving, ontology-based KM solution for monitoring changes in domains relevant for technology intensive industries.

In the following subsections we explain the core components of the KMP in further detail. Thereby we describe the architectural design of the components as applications of the NGKM-technologies. The major contribution referred to in the following is the combination of the NGKM-enabling technologies presented in section 3 into a real-world KMS which had not been achieved adequately at the design time of the KMP.

4.2.1 Dynamic Ontology Management System

As depicted in Figure 2, the Dynamic Ontology Management System DOMS is the place for storing and maintaining the ontologies. The ontologies are used for knowledge representation throughout the whole KMP. First the initial ontologies are deployed here, secondly the personalized knowledge maps for users are derived out of the ontologies, thirdly the Search Module Engine takes ontology data as an input, forth the pre-processed search results are stored as ontology instance data in the repository, and fifth the ontologies are updated according to changes in the respective knowledge domain. This extensive usage of ontologies in the KMP points out the importance of ontologies as the underlying data model in the KMP. Thus secure data storage and retrieval system as well as reasonable ontology maintenance techniques are required. We describe the technical realization for ontology storage and maintenance in KMP and then show how this solution supports NGKM.

As ontology storage device Sesame has been chosen.[4] Sesame is an open source storage and querying system for RDF-data which has been designed and implemented to be independent of any storage devices. As such, it can be implemented on top of RDBMS, OODBMS, and triple stores without having to change the query Sesame interacts with the DBMS through the Repository Abstraction Layer (RAL) interface. RAL communicates with the three functional modules of Sesame, namely: the RDF

[3] RDF: Resource Description Framework. Modelling framework for semantic description of information, especially web-documents. An XML-Serialization enables RDF-data exchange over the internet. RDF is a basic ontology specification technique for the Semantic Web. RDF is a W3C Recommendation (Lassila and Swick, 1999).

[4] Sesame was developed in the On-To-Knowledge project (Davis et al., 2002)., see http://sesame.aidministrator.nl/ It is considered as one of the most mature solutions for storage and querying of RDF-data (Magkanaraki et al., 2002).

Administration Module which enables the insertion and deletion of data, the RQL Query Module which enables querying of ontology instance data and the RDF export module that exports schema and instance data in RDF-format. These modules then communicate with their clients either via HTTP or SOAP.

Sesame provides a secure and scalable environment for storing and retrieving RDF-data, and its applicability has been extensively discussed (Davis et al., 2002). A further benefit of Sesame is that there is an additional module available, the Ontology Middleware Module (OMM), which provides further ontology management facilities required in the DOMS. OMM is an extension of Sesame which adds a versioning and a security mechanism to Sesame (see Figure 3). The reasons why OMM provides the ontology management facilities needed in the DOMS are discussed below.

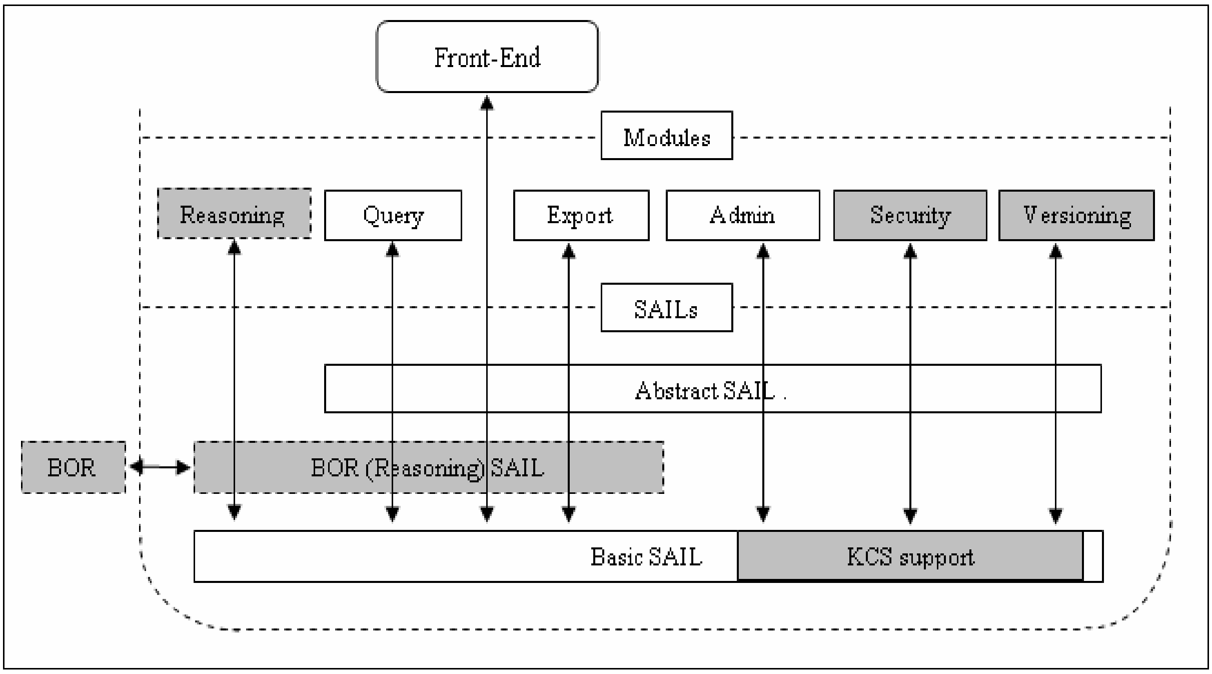

Figure 3: Sesame + OMM Architecture[5]

To allow updating of ontology schema and instance data, the repository has to provide a mechanism for ontology versioning. This means that ontology schemas can be updated in the repository and for each change a new version is defined. The problems arising when ontology data are changed are (i) that the ontology might gets inconsistent (i.e. a concept is deleted which an existing instance belongs to) and (ii) that a change may affect the functionality of applications or system components that apply the ontology as their data model. The general approach for ontology versioning thus distinguishes two kinds of changes in an ontology (Kiryakov et al., 2002b):

1) **Addition of a statement:** adding a new concept or an instance to the ontology. Changes of this type do not affect the information consistency of the ontology or the functionality of system depending on it.
2) **Removal of a statement:** deleting a concept or an instance. Removing a concept might affect the ontology's consistency (the semantic correctness needs to be controlled manually, see below). Any removal might affect the functionality of

[5] Source (Kiryakov et al., 2002a). The Reasoning Module of OMM is not utilized in the h-TechSight DOMS.

depending systems. These affects have to be detected and handled by the versioning system. A modification is considered as a combination of Addition and Removal.

The OMM-versioning mechanism assigns a repository an Update Counter (UC), an integer which is increased every time the repository is updated. Version Information is added to every RDF-statement (schema data & instances) called an Update Identifier UID. The UID contains information at which UC-number the RDF-statement has been added or deleted. The current state of the repository (= current ontology version) is constitutes by the "living" RDF-statements (Kiryakov et al., 2002a). This way of ontology versioning is very effective since only those ontology items are changed that are actual updated. The consistency of the ontology data in the DOMS is ensured by performing ontology updates manually (see below for further explanation). To assure that existing instances are assigned to corresponding concepts after a concept modification, the ontology assignment tool implemented in the Knowledge Discovery Module (see section 4.2.3) is invoked after an ontology update. Since all components of the KMP retrieve their data directly from the repository, ontology updates can not influence its functionality.

The second feature of OMM used in the DOMS is a security mechanism which allows defining access restrictions to the ontology repository. This is used for defining different user roles in the KMP. The distinction of different user roles for the KMP has several reasons: first the consistency and correctness of ontology schemas is important as they provide the underlying data model of the whole system. Therefore write-permission to ontology schemas should be limited to a minimal extent. Secondly information protection has been detected as a user requirement of high importance. Since the OMM Security Module allows defining access permission on the data level, information protection in the KMP is implemented via user roles. The following user roles are defined in the KMP: ´Ontology Managers´ control the evolution of an ontology, perform updates and thus have full access (several ´Ontology Managers can be named in the KMP). ´Users´ generally have reading-rights for the repository. In order to achieve information protection, these can be further restricted to certain ontologies or specific ontology data. The distinction of user roles in the KMP applies Role Theory as an important concept in KMS design by defining different positions with distinct required abilities and lines of action (Biddle, 1979).

The last functional aspect compulsory for enabling ontology updating is a method for ontology integration, i.e. the combination of ontologies. In KMP this is needed when changes in a knowledge domain occur and the corresponding ontology schema in the repository has to be updated. Recent research approaches in the field of ontology integration provide semi-automatic tool support. The most promising solutions are PROMPT[6] and Chimaera[7]; these tools support ontology integration by evaluating the ontology structure along with enhanced linguistic features, but the final integration decision is left to human intervention. In the KMP, this task is performed manually as it requires a sophisticated familiarity with ontology techniques. As this cannot and should not be expected of all users of the KMP, Ontology Managers (see above) – which are supposed to be the especially skilled for this – will perform ontology integration task within the KMP.

[6] PROMPT: PlugIn for the PROTÉGÉ – editor. See: http://protege.stanford.edu/plugins.html

[7] See: http://www.ksl.Stanford.EDU/software/chimaera/

The ontology storage and maintenance techniques used in the DOMS realize a convenient solution for storing, retrieving and maintaining ontology data in KMP according to the system's requirements. By this, the DOMS enables the utilization of ontologies as the grounding data model in KMP and thus incorporates the benefits of ontology techniques for NGKM pointed out in section 3.1

4.2.2 Search Module

The Search Module performs the actual search and monitoring of web resources. As stated above, the KMP distinguishes two different search modes: the so-called generic search mode, realized by the MAS Search Engine which searches the internet for a search request, and the application search mode which allows monitoring specific web resources via discrete queries. In the following we describe the functionality of both search modes.

The MAS provides an ontology-based web search via agents which allow automating search request performing. As the architecture and workflow of the MAS has already been described elsewhere (Aldea et al., 2003) we will only summarize this briefly in order to point out the contribution of the MAS to NGKM-technology. A user specifies a (sub)ontology that he wants to receive information for from the Internet. The ontology schema data is passed to the MAS as RDF-data. The MAS takes the ontology concepts as input and performs a search on the Internet. Then the workflow for the web-search is as follows:

(1) Coordinator Agent: decomposes ontology sub tree to be searched for into "query ontologies" (1 concept = 1 query ontology).

(2) Internet Agent: search the web for each query ontology (= term / concept). Therefore, recent web search engines are used (currently APIs for AltaVista and Google). The number of Internet-Agents can be adapted to the number of ontology concepts to be searched for, thus a reasonable scalability of the MAS can be achieved.

(3) Coordinator Agent: (re)composes the search results to taxonomic structure of the ontology sub tree searched for (the opposite of the decomposition in step (2)). The search results are sent to the Knowledge Discovery Module for further processing.

The MAS realizes NGKM-requirements in several ways. First, the web search is based on ontologies whereby the precision of retrieved information can be increased tremendously because the amount of irrelevant search results is reduced (see section 3.1 for detailed discussion of ontology-enhanced web search). Secondly the web search, i.e. information update for KMP users, can be performed automatically since the MAS uses agents. Thirdly the output of the MAS is semantically described web resources: the search results are related to the ontology concepts and thus comprise ontological notions which can be applied by the subsequently knowledge discovery techniques. Last but not least, the MAS presents a generic solution for ontology-based web search since every kind of ontology can be used as an input.

The application mode is realized as by WebQL-scripts. WebQL is a system for querying internet resources, consisting of a query language along with a query engine. It allows querying of conventional websites as well as the content of data repositories accessible of the web in the style of SQL for RDBs.[8] In order to define a suitable WebQL-script,

[8] WebQL: commercial tool developed by QL2 Software. see www.web.com

the internal structure of the respective web resource has to be known a priori which limits the usability of WebQL to a small set of web resources known by the system designer (Kuhlis and Tredwell, 2003). Nevertheless, it allows to automatically monitor and extract information from specific web resources detected as information sources of long-term interest, for example the website of a journal or a relevant web portal, in a much more sophisticated way than conventional methods for observing internet content.

4.2.3 Knowledge Discovery Module

The Knowledge Discovery Module is the third building block of the KMP. It implements facilities for analyzing the web search and web mining results, for assigning information as instances to the ontologies in the DOMS, and for updating ontology schemas. In the following we outline the conceptual ideas of the knowledge discovery techniques applied, point out how the Knowledge Discovery Module completes automated and cyclic knowledge item processing, and in which way NGKM-enabling technologies are applied.

As depicted in Figure 2, the Knowledge Discovery Module implements three kinds of features. First, a Content Analysis which investigates the information contained in the search results, secondly, a facility to assign information to the existing ontologies as instance data, and thirdly, a feature for discovering conceptual changes in a knowledge domain in order to propose an ontology update. We explain these three features in more detail.

Content Analysis

The search results of the MAS are links to web-sites that contain relevant information for the search term, i.e. the (sub)ontology that has been searched for. The information contained in these web-resources are analysed – that means the actual content of the web-sites is investigated. Therefore existing tools are customized which perform text mining on the web documents by recent NLP-techniques (tokenization, POS-Tagging, Structural Filtering) in order to achieve extraction of terms that correspond to the ontology schema definitions. The output of this sub-module constitutes the input for the following features. The same is done with the information retrieved from the Web Mining tool.

The Content Analysis in KMP relies for the most part on the GATE system developed at the University of Sheffield. The GATE system is a tool for ontology-based information extraction which supplies a collection of pre-defined NLP-based tools and its usability has been proven in an ostensive number of applications (Cunningham et al., 2002). According to the NGKM-enabling technologies introduced above, the features implements NLP-techniques as the foundation for automated web content processing (see section 3.2).

Ontology Instance Assignment

The second feature is the alignment of the web documents retrieved and analyzed as instances to the ontologies in the DOMS. This facility is implemented in the so-called 'Ontology Enrichment Module'. On basis of the terms extracted by the Content Analysis, the web resources are assigned a weight-value which determines the rate of conformity of the document content and an ontology concept. Further, the web documents are clustered in accordance to an ontology concept and then assigned as instances to the concept ordered by their weight-values. Thereby a continuous evolvement of the KMP knowledge base, i.e. the information available in the DOMS, is enabled. Moreover, the instances of an ontology concept will be ordered according to their ontological affiliation whereby a surplus for users in terms of quality of retrieved information is achieved and – as the Search Module and the Knowledge Discovery Module can be invoked automatically – the knowledge base will contain the most recent information for a knowledge domain.

This feature is implemented mainly on basis of the THESUS tool developed at the Athens University of Economics and Business (Varlamis et al., to appear). The Ontology Instance Assignment facility is to be classified as a KD technique as introduced in section 3.3.

Change Detection in Knowledge Structures for Ontology Updates

The third functional facet provided in the Knowledge Discovery Module is the detection of changes in the knowledge structure of a domain in order to update the corresponding ontology schema. This feature relies on the same technologies that are applied for Ontology Instance Assignment. The weight-values and the semantic clusters computed for the web documents are compared to the ontological structure of the existing ontologies in the DOMS. If there are significant differences between the keyword-cluster (which represents the ´real´ content of a web-document) and the ontological structure, guidelines for changes to the ontology schema are derived automatically. For the reasons stated above the updating of ontologies, i.e. the integration of these ontology update proposals with an existing ontology is performed manually by an Ontology Manager. According to the classification of NGKM-enabling technologies introduced, this feature realizes a KD technique.

Summarizing, the Content Analysis improves the quality of information stored in the KMP via NLP-techniques. The automated Ontology Assignment allows dynamical evolvement of the information, i.e. the ontology data available in the system, and the Change Detection ensures that the semantic knowledge structure is kept up to date. Thus the Knowledge Discovery Module is a very important contribution of the h-TechSight KMP to NGKM-technology since the three features described enable automated knowledge item processing which has been outlined as the most important challenge for NGKM-technologies.

4.3 Project Outcome and Research Results

After discussing the technological realization of the KMP thoroughly, we will now attend the outcome of the h-TechSight project concerning the design and the functional quality of KMP. In particular, we outline the experiences assembled in designing the KMP and position it as a next generation KMS.

From a functional point of view, the KMP is designed as a system for automatically monitoring and assessing changes in information resources which users define to be relevant, especially concerning the external environment of organizations. Thus, it does not support the complete intra-organizational KM-process pictured in Table 1, but provides a functional component which can be integrated into existing KM-systems or be used as a starting point to build up an individual KM- solution for an organization.

4.3.1 Findings on KMS Design

In the following we explain motivations, decisions, and experiences assembled during the design of the KMP. We do not want to claim these insights as general design guidelines, but we consider them as valuable information for building profitable KMS.

The first, possibly the most important aspect to mention is the componentized design of the KMP functionality. This does not refer to the component-orientated technical architecture, but to the fact that the KMP serves mainly as a tool for automated acquisition of relevant information and thus provides a modular functionality instead of a generic tool suite to support intra-organizational KM within all its facets. The offered functionality is of general interest for KM and can be used in different ways for the individual of an organization. In the h-TechSight project consortium, for instance, there are large-scale companies that already run a KMS and integrate the KMP as a new components into their system, while in some SMEs which currently have no or only a very immature KM-solution the KMP is used as the central application whereupon the KM-system for knowledge dissemination and application is built. This is considered a significant advantage by the end-user-group in the project consortium because the usability of the KMP does not presume specific organizational pre-requisites, thus it can be used in a wide variety of application scenarios. As the KMP does not offer a complex tool suite, it is easy to handle and not very cost-intensive which makes it affordable for SMEs with limited KM-budgets.

A second aspect to be examined in more detail is the usage of ontologies as the grounding data model and knowledge representation technique in the KMP. Apart from the technical benefits for information processing mentioned, the following functional advantages of ontologies are considered to be most important. First, an ontology as a conceptual model of a domain allows reducing complexity by structuring information which is a very significant support for identifying relevant aspects for KM. The KMP-feature for personalizing ontologies to specific user needs enables re-use of existing ontologies, thus facilitating this process. A second functional benefit of the KMP is the evolvability of ontologies which ensures keeping the information in the system up-to-date, whereby the ontology versioning mechanism allows automatically maintaining the information repository. Thirdly, ontologies that are accepted as common models of a domain, like the sample ontologies provided by the h-TechSight consortium, ease communication and information exchange in a community (either inside an organization or between organizations) as an essential support for KM. In conclusion, the experiences of the h-TechSight project in using ontologies as the basic KM-technique are very positive and strengthen the current trend of ontology-based KM-solutions.

As a last aspect concerning the design of KMS we want to stress out that the design of the KMP has been a very complicated and long-term process. Although is basic idea of the system is straightforward and the objectives in project consortium were homogeneous, it was very challenging to determine on the actual features to be provided and to derive a common line of action. Due to this, we believe that identifying the functional objectives in detail is crucial for creating profitable KMS. Although this might valid for IT-systems in general, a close cooperation between the end-users and the technology providers seems to be an essential premise for the development of KMS.

4.3.2 Positioning of h-TechSight KMP as a NGKM-solution

In order to provide a structured description of the KMP functionality and to position it as a NGKM-solution we expose in what way the h-TechSight KMP accomplishes the requirements for NGKM-solutions depicted in section 2.2.

Supply – Demand Side KM

The KMP offers a combination of supply- and demand-side KM. The sample ontologies provided supply an initial knowledge structure which can be customized by each user. Via personalized ontologies the user retrieves up-to-date information available on the Internet for his individual needs as demand-orientated KM. The updatability of ontologies allows sharing evolvements in the knowledge structure of a domain throughout the user community which supports supply-side KM driven by the user community.

Knowledge Life Cycle

The functional objective of the KMP is to provide a tool for automatically monitoring internet resources for information relevant to the external environment of an organization, not to support the complete internal knowledge life cycle of an organization. Tracing changes in external knowledge sources is supported by evolving ontologies in the KMP which allows keeping information in the system, and thus the knowledge of an organization, up to date.

Knowledge Processes

The support of knowledge processes inside an organization – which this aspect refers to – is not a central aim of the KMP. Nevertheless, the KMP provides an information repository with sophisticated information browsing and retrieval functionalities which can easily be integrated into an organizational knowledge process.

Knowledge as Rules

In its definition, this aspect also refers to internal KM of an organization. Seen from a broader point of view, not only the knowledge inside an organization determines the action of the organization but also the knowledge in its external environment. The main functionality of the KMP is to monitor and assess internet resources relevant to the external environment of an organization, thus it helps to discover knowledge which rules the organization's actions.

Knowledge Structures

Due to its nature, an ontology defines knowledge structures as a conceptual model of a certain domain. Since ontologies are applied as the underlying knowledge representation model in the KMP, the NGKM-requirement of enhanced techniques for capturing and transferring knowledge structures is realized to a very high extent.

Nested Knowledge Domains

This aspect requires support for knowledge communication between individuals, groups and an organization as a whole. The h-TechSight KMP assists this knowledge sharing since the KMP can be used by individuals and groups; besides, the evolving ontologies provide an updatable, commonly shared understanding of a domain which facilitates knowledge communication. Moreover, defining cross-domain ontologies (e.g. general employment ontology) allows nesting of different industrial sectors.

Organizational Learning

As pointed out above, the evolvability of ontologies as the grounding knowledge model of the KMP allows users to monitor changes in their field of interest. By this, relevant changes in quickly and dynamically evolving external environment can be detected as crucial information for organizational learning in order to keep competitive.

Complexity Theory

As pointed out in section 3.1, ontologies are a suitable means for complexity reduction since they allow representation of knowledge structures in a highly expressive manner and they further allow splitting complex domains into smaller, better comprehensible parts. Moreover, ontologies represent a conceptual model for a domain which decreases the complexity for human understanding for the domain. As the KMP employs ontologies as its grounding technique for knowledge representation, it realizes these advantages.

5. RELATED WORK

Work related to the h-TechSight project and the development of the KMP is general design of KMS on the one hand and research activities on NGKM-enabling technologies on the other. We will depict both aspects and relate current efforts to the work presented in this paper.

5.1 KMS Design and Development

Development of KM-solutions and the design of KMS has spawned a huge amount of approaches and literature. Although several aspects related to KM and the development of KMS have been studied, terminology and design guidelines approaches are still ambiguous. We will stick to the terminology used throughout this paper for positioning the KMP in a framework for structuring KMS, point out related theoretical approaches and briefly discuss a KM-solution similar to the KMP.

Hahn and Subramani present a framework to characterize KMS and point out challenges arising in KMS design (Hahn and Subramani, 2000). For classifying KMS, a two-dimensional matrix is defined differentiating the type of knowledge processable within a KMS and the level of a-priori structure of knowledge. In this framework, the KMP is to be allocated as a KMS for collaborative filtering of unstructured knowledge artefacts on the internet. Besides, the framework identifies particular challenges in KMS design: at first, the balance between information overload and potentially useful information which is satisfied by the KMP because it monitors all possibly relevant internet resources and only retrieves information that correspond to the search request by means of ontology-enhanced web search and web mining. Secondly, the additional workload implied by a KMS should be in due proportion to the added value for KM-activities. In the KMP, additional work is only required for defining or selecting appropriate ontologies while system handling during runtime is equal to convenient web-based applications. Another aspect mentioned is the acceptance and usage of a KMS in an organization which is addressed in the h-TechSight project by explicit tasks assigned to the end-user groups in the project consortium.

A design theory for KMS is provided in (Markus et al., 2002) which mainly identifies six design principles. Although the model focuses on IT-system design for supporting internal knowledge processes in an organization, the design principles can be generalized for KMS design. Summarizing, the theory states that for development of a profitable IT-system to support KM in an organization, there should be a close, ´dialectical´ collaboration between users and the technology providers, the system should be designed for naïve users and allow knowledge sharing, and the functionality of the system should be componentized. While the former aspects have been followed throughout the development process of the KMP, the componentization-principle is an essential characteristic of the KMP functionality as mentioned above.

A KMS comparable to the KMP in terms of functionality and used technologies is the Volvo Information Portal (VIP) which extends the existing competence management system at Volvo IT (Lindgren and Stenmark, 2002). VIP is an agent-based recommender system which allows automatically searching the Volvo IT intranet for relevant information according to the user profile, similar to the MAS Search Engine of the KMP. Further, the systems implements a feature for finding persons with similar interests, either by matching user profiles or by matching agent signatures, thereby enhancing competence management and provoking communication among organization members. In contrast to the KMP, the VIP system does not use ontologies to describe information structures. On the other hand, the KMP does not yet provide explicit community features and communication or information exchange in a community of interest. Such functionality can be added to the KMP and is considered for future development.

5.2 Related Research on NGKM-solutions

Currently, many research efforts are engaged with NGKM-enabling technologies. We will shortly describe two IST-5th framework projects which have already been finalized and are frequently referred to as grounding initiatives for Next Generation KM. There are a lot of other initiatives concerned with NGKM-solutions which will not be enlisted here in detail.[9]

On-To-Knowledge[10]

The objective of the On-To-Knowledge project was to create a methodology and a tool suite for ontology-empowered, web based knowledge management. The application area in focus was ontology-based web content processing, i.e. ontology extraction from web documents, linking ontologies to huge amounts of data and enable ontology data storage and querying.

The tool suite developed in the project comprises tools for ontology generation out of structured and unstructured documents, solutions for storage and querying of ontology data (among other Sesame and OMM which are also utilized in the h-TechSight KMP, see section 4.2.1), and several tools for ontology-based information presentation and web-search.. Further information can be found in (Davis et al., 2002) or on the project website.

The major contribution of On-To-Knowledge to NGKM research is the development of the tool suite that enables ontology-based information processing with special regards to web-orientated Knowledge Management. As this tool suite has been the first one

[9] See “Knowledge Board” for further information, http://www.knowledgeboard.com/index.html

[10] On-To-Knowledge: Content-driven Knowledge Management Tools trough Evolving Ontologies. IST project (IST-1999-10132). Website: http://www.ontoknowledge.org

providing tools for all aspects of handling ontology data it is considered as the starting point for ontology-based KM and has been used as an important reference in the h-TechSight project. Another important contribution of On-To-Knowledge is a KM-Methodology specialized on ontology-based KM and embraces the ontology development process as a specialisation of the KM-activities 1) – 4) of the general KM-model presented in Table 1 (Sure et al., 2002b).

VISION[11]

The VISION project aimed at defining a scientific and a technological roadmap for NGKM-development. The projected reviewed the state-of-the-art in NGKM-research (concerning technology as well as non-technological aspects) and derived research requirements for NGKM to be exploited in a project to follow up the VISION project.

The working areas of the VISION project have been to review existing show cases in the context of organisational knowledge management and the state of the art in NGKM-enabling technologies like Semantic Web technology, knowledge discovery, processes and groupware, and mobile technologies, and to build bridges between the fields. A roadmap for NGKM as a basis for future research and development was designed. A further aim was to build up a network of organizations and institutions that are engaged with NGKM and to provide a web portal for information resources on NGKM (Software, Organizations, Projects, Topics and Persons concerned with NGKM).

The most important results with regard to the relevancy on NGKM research are surveys on the state-of-the-art on NGKM-technologies (Anagnostou et al., 2003), a "Strategic Roadmap" that analysis the maturity of NGKM-technologies (Abecker et al., 2003), and the Maturity Model that complies critical success factors of NGKM Methodologies (Weerdmeester et al., 2003).

In contrast to recent research efforts in NGKM, especially the NGKM-projects introduced above, the KMP does not offer completely new developed KM-tools but applies existing technologies and combines them in order to provide a NGKM-application for a specific user community. Since there are suitable technological components existing for the processing features needed in the KMP, there was no need to develop new tools to achieve the objectives of the h-TechSight project. The technological challenge persecuted in the KMP development has been to combine technologies in to a profitable real-world KMS.

6. CONCLUSIONS

In this paper we have discussed recent trends and requirements concerning the functional design and technological realization of the next generation of Knowledge Management Systems. On basis of this, we have presented the Knowledge Management Platform KMP developed in the h-TechSight project as a NGKM-solution. The KMP allows automatically monitoring and accessing internet resources a user defines to be of interest, especially intended to facilitate monitoring the external environment of an

[11] VISION Project Website: http://km.aifb.uni-karlsruhe.de/fzi/vision/

organization. We described the technical realization of the KMP and pointed out experiences and research results assembled on KMS design.

The Next Generation of Knowledge Management consolidates functional and technical requirements on KM and particularly on KMS in order to overcome the limitations of recent KM-solutions. Concerning functionality, next generation KMS should provide better support for intra- and inter-organizational KM-activities by more sophisticated assistance for demand orientated KM, more accurate techniques for supporting the organizational knowledge lifecycle and representing knowledge structures, and facilitate organizational learning by enhanced means for knowledge identification, creation and allocation. This imposes requirements for enabling technologies; although knowledge is always related to humans, KM-technologies are the building blocks of KM-solutions and they determine the quality of the KM-support provided the complete system. As conventional technologies employed as the technological basis of KMS have significant restrictions for automated knowledge-item processing, other, AI-based technologies are considered as the enabling technologies for NGKM-solution. We have investigated three groups of technologies and pointed out why they are considered as NGKM-enabling technologies: (1) Ontologies as highly expressive knowledge modelling and representation technique which can be used especially to enhance the quality of web search, (2) Natural Language Processing technologies for automated web content analysis, and (3) Knowledge Discovery technologies that enable automated knowledge item processing. These groups are very general and can be realized by different technologies, but they provide the technological features required for NGKM-solutions.

The KMP developed in the h-TechSight project is designed for automatically monitoring and reporting changes in the external environment of an organization by using the Internet as its information source. The KMP supports structuring user information needs into ontologies, supplies up-to-date information from the Internet according to individual user requirements, and provides automated updating and pre-processing of this information. This is realized by the three building blocks of KMP. The DOMS enables secure storage and updatability of ontologies as the grounding data model, the Search Module comprised of an agent-based general search facility and a web mining tool for monitoring distinct web resources allows ontology-enhanced information retrieval from the web, and the Knowledge Discovery Module realizes automated knowledge item processing based on NLP- and Knowledge Discovery Techniques. It is important to note that the h-TechSight KMP provides a modular solution for monitoring changes in dynamically changing external environments employable as a component for a KMS individually designed according to the actual user needs.

In the KMP development process, several experiences and decision have been made which might serve as insights for future development of next generation KMS. In particular, the componentization-principle for the functional design of a KMS, i.e. to provide a modular functional feature that is of general need in KM like the KMP functionality, has been approved as very valuable because it does not impose certain conditions on the organization structure for using the system and thus broadens the applicability of a KMS. Besides, such modular systems are not very cost-intensive which makes them affordable for SMEs. A further outcome of the h-TechSight project is an affirmation on the usage of ontologies as the underlying KM-technique. Ontologies do not only offer technological benefits for enhanced information processing, but also serve as a sophisticated means for presenting knowledge structures, reducing complexity and facilitate information exchange and communication in a community. Last but not least, it was detected that the design of a KMS is a very complicated task which requires a close and iterative collaboration between the users and the technology providers as performed in the three-phase development process of the KMP. Even more than in development of conventional IT-systems, the detailed definition of system requirements and functionality seems to be crucial for developing profitable KMS.

Concluding, we summarise that the technical accomplishments and the insights on functional design of KMS obtained in h-TechSight project provide a contribution to the development of next generation KMS with regard to the requirements disposed for NGKM-solutions.

Acknowledgements

This research has been partly funded by the h-TechSight project, IST-2001-33174. We would like to thank the h-TechSight consortium as well as the members of DERI – Digital Enterprise Research Institute for fruitful input to this work.

References

Abdecker, A., van Elst, V. (2003a). Ontologies in Knowledge Management. In S. Staab, R. Studer (Ed.): *Handbook on Ontologies.* Series in International Handbooks on Information Systems. Berlin, Heidelberg: Springer, pp. 435-454.

Abecker, A.; Hefke, M.; Maedche, A.(ed.). (2003b). *Final Version of Strategic Roadmap*. VISION Deliverable D4.2, VISION Consortium. Available at: http://km.aifb.uni-karlsruhe.de/fzi/vision/vision/docs/D4.2_final.pdf

Alavi, M.; Leidner. D. E. (1999). *Knowledge management Systems: Issues, Challenges, and Benefits.* Communications of the AIS, 1(7) : 1-37.

Alavi, M.; Leidner, D. E. (2001). *Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues*, MIS Quarterly (25:1), pp. 107-136.

Aldea, A. et al. (2003). An Ontology-Based Knowledge Management Platform. In: Proceedings of the Workshop on Information Integration on the Web (IIWeb-03) at the *18th International Joint Conference on Artificial Intelligence*. Acapulco, Mexico.

Allen, J. (1994). *Natural Language understanding*. 2nd edition, Redwood City, CA (USA): Benjamin/Cunnings.

Anagnostou, A.; Panagopoulou, G.; Patra, E.. (2003). *Final Version of State of the Art Report on Core and Extended Enabling Technologies.* VISION Deliverable D3.2. Available at: http://km.aifb.uni-karlsruhe.de/fzi/vision/vision/docs/D3.2.pdf

Applehans, W.; Globe, A.; Laugero, G. (1998). *Managing Knowledge. A Practical Web-Based Approach.* Boston: Addison-Wesley.

Bhatt, G. D. (2001): *Knowledge Management in organizations: examining the interaction between technologies, techniques , and people.* Journal of Knowledge Management 5(1), pp. 68-75.

Biddle, B. J. (1979). *Role Theory. Expectations, Identities and Behaviours.* Academic Press: New York.

Braa, K.; Vidgen, R. (1999). *Interpretation, intervention, and reduction in the organizational laboratory: A framework for in-context information systems research*, Accounting Management and Information Technologies, Vol. 9, 1999, pp. 25-47.

Cunningham, H.; Maynard, D.; Bontcheva, K. and Tablan, V. (2002). A framework and graphical development environment for robust NLP tools and applications. Proceedings *of 40th Annual Meeting of the Association for Computational Linguistics (ACL2002)*, pp.168-175. Philadelphia, PA, USA, July 6-12.

Davenport, T.H.; Prusak, L.(1998). *Working Knowledge. How organisations manage what they know.* Boston, USA: Havard Business Scholl Press.

Davis, J.; Fensel, D.; van Harmelen, F. (ed.). (2002). *Towards the Semantic Web. Ontology-Driven Knowledge Management.* London: Wiley.

Engels, R.; Bremdal, B. (2000). *Information Extraction: State-of-the-Art Report.* On-To-Knowledge Deliverable D5.

Fayyad, U.; Piatetsky-Shapiro, G.; Smyth, P.; Uthurusamy, R.(Ed.) (1996). *Advances in Knowledge Discovery and Data Mining*. Boston: AAAI/MIT Press.

Fensel, D. (2003). *Ontologies. A Silver Bullet for Knowledge Management and E-Commerce*. 2nd Edition. Berlin, Heidelberg: Springer.

Fernández, M. (1999). Overview of Methodologies for Building Ontologies. *Workshop on Ontologies and Problem-Solving Methods: Lessons Learned and Future Trends (IJCAI99)*.

Fernando, D et al. (2002): *Requirements for KM in technology intensive industries.* h-TechSight Deliverable D1. Available at: http://prise-serv.cpe.surrey.ac.uk/techsight/seefiles.asp?delid=8

Fernando, D.; Kokossis A.; (2003). *H-TechSight boosting competitive performance*, The Chemical Engineer, September 2003.

Gómez-Pérez and O. Corcho. (2002) Ontology languages for the semantic web. *IEEE Intelligent Systems*, January / February, 54-60.

Gruber, T. R. (1993). *A translation approach to portable ontologies.* Knowledge Acquisition, 5(2):199-220.

Hahn, J.; Subramani, M. R.(2000). A Framework of Knowledge Management Systems: Issues and Challenges for Theory and Practice," in *Proceedings of the International Conference on Information Systems 2000*, pp. 302-312.

Kalakota, R; Robinson. (2000). *M.: e-business 2.0. Roadmap for Success.* New York, London: Addison-Wesley.

Kosala, R.; Blockeel, H. (2000). *Web mining research: A survey*, SIGKDD Explorations, 2(1), pages 1 - 15.

Kiryakov, A.; Ognyanov, D.; Popov, B. (2002a). *Ontology Middleware. System Documentation. Overview, Installation, Reference Guide.* OntoText, Bulgaria.

Kiryakov, A.; Simov, K.; Ognyanov, D. (2002b). Ontology Middleware and Reasoning. In J. Davis, D. Fensel, F. van Harmelen (Ed.): *Towards the Semantic Web. Ontology-Driven Knowledge Management.* London: Wiley, pp. 179-196.

Krohn, U.; Davis, J. (2001). *The Search Facility RDFferret.* On-To-Knowledge Deliverable 11, On-To-Knowledge Consortium.

Kuhlis, S.; Tredwell, R. (2003). Toolkits for Generating Wrappers - A Survey of Software Toolkits for Automated Data Extraction from Websites, in: M. Aksit, M. Mezini, R. Unland (Ed.): *Objects, Components, Architectures, Services, and Applications for a Networked World*, LNCS Vol- 2591, Springer: Berlin, S. 184-198.

Lassila, O.; Swick, R. R. (1999). Resource Description Framework (RDF) Model and Syntax Specification. W3C Recommendation.

Lindgren, R.; Stenmark, D. (2002). *Designing Competence Systems: Towards Interest-Activated Technology*. Scandinavian Journal of Information Systems 14 (1), pp. 19-35.

Maedche, A.; Motik, B.; Stojanovic, L.; Studer, R.; Volz, R (2003). *Ontologies for Enterprise Knowledge Management*, IEEE Intelligent Systems 18(2), pp26-33.

Magkanaraki, A.; Karvounarakis, G.; Anh, T. T.; Christophides, V.; Plexousakis, D. (2002): *Ontology Storage and Querying.* Information Systems Laboratory, Foundation for Research and Technology Hellas: Technical Report 308.

Markus, M. L.; Majchrzak, A.; Gasser, L. (2002). *A Design Theory for Systems That Support Emergent Knowledge Processes*, MIS Quartely 26 (3), pp. 179-212.

McElroy M. W. (2002) *The New Knowledge Management. Complexity, Learning, and Sustainable Innovation*. Boston: KMCI Press, Butterworth-Heinemann.

Milton, N., Shadbolt, N., Cottam, H. and Hammersley, M.(1999). *Towards a Knowledge Technology for Knowledge Management*, Human-Computer Studies 51(3), pp. 615-664.

Nickell, S. (1995). *The Performance of Companies: The Relationship between the External Environment, Management Strategies, and Corporate Performance*. Oxford: Blackwell Publishers.

Nonaka, I.; Takeuchi, H. (1995). *The knowledge creating company.* Oxford University Press, Oxford.

Polanyi, M (1967). *The Tacit Dimension.* London: Routledge and Keoan Paul.

Schreiber, G. et al. (2000). *Knowledge Engineering and Management. The CommonKADS Methodology.* Boston: MIT Press.

Smith, R.G.; Farquhar, A. (2001). *The Road Ahead for Knowledge Management: An AI Perspective*, AI Magazine 21(4), 2000, pp. 17-41.

Staab, S. et al. (2000). AI for the Web - Ontology-based Community Web Portals. *17th National Conference on Artificial Intelligence and 12th Innovative Applications of Artificial Intelligence Conference (AAAI 2000/IAAI 2000),* Menlo Park/CA, Cambridge/MA: AAAI Press/MIT Press.

Stenmark, D.(2001). *Leveraging Tacit Organisational Knowledge*, Journal of Management Information Systems 17(3), pp. 9-24.

Studer, R.; Benjamins, V.R.; Fensel, D. (1998). *Knowledge Engineering. Principles and Methods.* Data and Knowledge Engineering 25 (1-2) 1998.

Sure, Y.; Erdmann, M.; Angele, J.; Staab, S.; Studer, R.; Wenke, D. (2002a). OntoEdit: Collaborative Ontology Engineering for the Semantic Web", In Proceedings of the *First International Semantic Web Conference 2002 (ISWC 2002),* June 9-12 2002, Sardinia, Italy.

Sure, Y.; Studer, R. (2002b). A Methodology for Ontology-based Knowledge Management. In J. Davis, D. Fensel, F. van Harmelen (Ed.): *Towards the Semantic Web. Ontology-Driven Knowledge Management.* London: Wiley, pp. 179-196.

Russel, S.; Norvig, P. (2003). *Artificial Intelligence. A Modern Approach.* 2nd Edition. Pearson Education: Upper Saddle River, New Jersey (USA).

Varlamis, I.; Vazirgiannis, M.; Halkidi, M. and Nguyen, B. *THESUS, a closer view on web content management enhanced with link semantics*, IEEE Transactions on Knowledge and Data Engineering, To appear.

Weerdmeester, R.; Pocaterra, C.; Hefke, M. (2003). *Knowledge Management Maturity Model.* VISION Deliverable D5.2. Available at: http://km.aifb.uni-karlsruhe.de/fzi/vision/vision/docs/D5.2-KM-Final.pdf

Wilensky, U., Resnick, M. (1999). *Thinking in Levels: A Dynamic Systems Approach to Making Sense of the World.* Journal of Science Education and Technology, vol. 8, no. 1, pp. 3-19.