Proceedings of the 3rd Workshop on Learning Software Organizations - LSO, Germany, 2001. PRELIMINARY VERSION

Supporting Knowledge Management in University Software R&D Groups

Christiane Gresse von Wangenheim[1], Daniel Lichtnow[2], Aldo von Wangenheim[2]

Universidade do Vale do Itajaí[1]

Universidade Federal de Santa Catarina[2]

Abstract. One important success factor for Software Research and Development organizations is their ability to systematically manage knowledge and information. Based on our experiences, we examine the characteristics and the activities that need knowledge-based support in software R&D organizations in academic environments. A tailored KM solution is outlined considering knowledge content, process, organization and technical infrastructure required. The approach is currently being established and evaluated in the context of an international research project.

Keywords: Knowledge Management, Corporate Memory Management Systems, Experience Factory

1 Introduction

Software Research and Development (R&D) organizations aim at developing complex software products and services outstanding in terms of innovation and creativity. Typically, they are composed of multiple interacting communities, each possessing highly specialized knowledge in an environment characterized through rapid changes, shorter development cycles and increased quality demands. In this context, knowledge - the experience, insights, and practical know-how and skills that humans possess and which guide their decisions and actions - is an important asset that makes individual and organizational intelligent behaviour possible and is basic to innovation and creativity. Therefore, the ability to share and leverage knowledge across the a Software R&D organization has been recognized as important for the success of the organization.

---

1. *Author's address:* Universidade do Vale do Itajaí/CES VII, Rod. SC 407, Km 04,CEP 88122-000 São José/SC, Brazil. *Tel:* +55 (48) 281.1517. *Fax:* +55 (48) 281.1506. *E-mail:* gresse@sj.univali.br *WWW:* http://www.sj.univali.br/~gresse.

2. *Author's address:* Universidade Federal de Santa Catarina, Trindade, 88049-200 Florianópolis, Brazil. *Tel:* +55 (48) 331.7552. *Fax:* +55 (48) 331.9770. *E-mail:* {lichtnow,awangenh}@inf.ufsc.br *WWW:* http://www.inf.ufsc.br/~awangenh.

In this context, Knowledge Management (KM) [7],[23] and Organizational Learning (OL) [8],[16],[21] have emerged as explicit areas of pursuit for managing organizations, focusing on how organizations can understand what they know or need to know and how to make maximum use of the knowledge, and also learning from its application, typically using advanced technology [7],[17].

KM improves knowledge circulation and communication in the organization supporting effective and efficient decision making and the creation of innovative and creative solutions, allowing an organization to retain critical expertise and to avoid loss of an expert's know-how after his retirement or mutation. KM enables the exploitation of the experience acquired from past projects and to keep lessons learned in order to avoid the repetition of mistakes and allows nonexperts to retrieve expert advice when needed. It reduces the duplication of effort by efficiently building on previous work. KM enables the consolidation of corporation-wide knowledge into competencies and shortens the learning curve for new technologies, empowering organizations to adapt quickly to changing opportunities [18].

In Software R&D organizations, KM is frequently done in an *ad-hoc*, informal manner, where the decision to reuse is made by individuals, and the type of knowledge reused is usually limited to personal experiences. Relevant knowledge is often not directly explicitly available, being only available implicitly in the heads of a few employees of the organization. Making this implicit knowledge explicit or sharing it is difficult. Another problem is to access valuable information when required and use it in an efficient and cost-effective manner due to information overload. To maximize productivity and quality gains through reuse of experience, KM has to be systematized.

This paper addresses the problems faced regarding KM in Software R&D organizations in academic environments, as current R&D methodologies do not adequately address and support the capture and use of relevant information and knowledge. Our work is applied in the context of an international research project, The *Cyclops* Project [6]. This R&D project aims at the development and transfer of new methods, techniques and tools in the area of Medical Image Analysis, being performed by an international R&D consortium consisting of the Federal University of Santa Catarina (Brazil), Taubaté University (Brazil), University of Kaiserslautern (Germany), GMD FIRST (Germany) and medical and industrial partners of both countries, representing an heterogeneous and geographically widespread environment.

The paper is structured as follows: in section 2, we examine the characteristics and specific problems encountered by academic R&D organizations and in Section 3, we describe the research and development activities for which we claim knowledge-based support ought to be provided. A KM solution customized to the identified characteristics, problems and needs is outlined in Section 4. Section 5 describes our current state of research and indicates future research directions.

2 Characteristics of University Software R&D Groups

The principal focus of academic R&D organizations in the Software domain is on research in computer science and the development of prototypical software systems. In this context, research means comprising creative work undertaken on a systematic basis in order to increase the stock of knowledge and the use of this stock of knowledge to devise new applications. It generally falls into two categories: applied research or basic research. Considering the expanding scope of computer science today, topics that are currently considered part of the computer science are technology rather than theory driven. This leads to the development of prototypical software systems which are implemented in order to demonstrate the research solution to be documented and shared. In the domain of our research project, the principal focus is on applied research in the areas of Intelligent Medical Image Analysis, Workflow Management in Medicine, Medical Image Databases and Teleradiology.

In this context, the main objectives of University R&D Groups are twofold:

- on the organizational level, they aim at developing complex software products and services outstanding in terms of innovation and creativity.
- on the individual level, they aim at supporting students to perform intellectually challenging academic projects as part of an honours degree.

This requires support for learning on both levels: organizational and individual, especially as one of the principal objectives of an university environment is to teach knowledge and capabilities.

R&D organizations are typically characterized through highly specialized knowledge and advanced technologies in specific research areas. Researchers need to have sound theoretical knowledge and practical experiences on the specific research topic as well as on research methodologies in order to effectively and efficiently perform research work and software development, as illustrated in Figure 1.

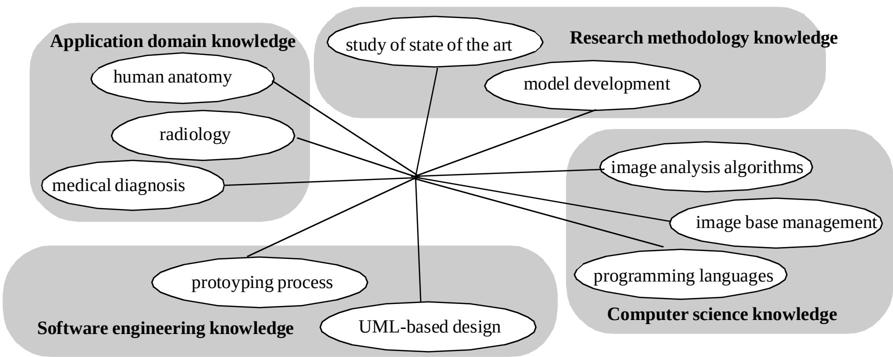

Figure 1 Examples of knowledge areas of interest

Regarding specifically software research and development in university environments, we can further observe the following characteristics:

Cross-functional and institutional collaboration: R&D organizations are typically composed of multiple interacting communities involving cross-functional and institutional linkages. Different participants join a research project with differing viewpoints. This includes, in the Cyclops research group, for example:

- Researchers in Computer Science
- Researchers in Medicine and medical doctors

- Representatives from radiological equipment and medical software companies.

This also shows the need to bring together participants from across multiple collaborating organizations, where, expertise and skills might be distributed both within and outside the organization. For example, for a research project in the Cyclops project, it is common to employ expertise in specialized areas such as radiology from medical clinics or, for example, through collaborating universities in Brazil and Germany. These collaborating organizations may be geographically distributed in Brazil or world-wide. Especially through the close interaction among these cross-institutional and functional teams, synergy effects contribute to research progress. However, such a collaboration brings in especially the need to organize, integrate, filter, condense and annotate relevant information that these team members contribute and to facilitate knowledge sharing and dissemination.

Constant turnover: University research groups are characterized through a constant turnover of participants, as students generally leave the group when finishing their Master or Ph.D. degrees. This is a major threat to the collective knowledge, since much of the knowledge is situated in the minds of individuals. If there is no repository for knowledge other than personnel, turnover leads to reduction in the organizational knowledge. This includes explicit knowledge (e.g., documents, reports etc. which have not been stored in a centralized DB) as well as tacit knowledge which has not been explicitly articulated and stored.

Knowledge leverage: In general, new group members which join the group have different levels of knowledge and capabilities, e.g., a new researcher in the Cyclops Project may have knowledge on image analysis but not on human anatomy or the programming language Smalltalk. In addition, new group members will have problems to get an overview on the research done in their area (including also the identification of experts) within or outside the organization. In order to enable the realization of effective research work, members have to learn the knowledge required for the execution of their research project. Such a workplace learning requires to be flexible so that it can be easily adapted to changes regarding, e.g., the research areas of interest.

Reinvention of solutions: Often it has been observed that, due to missing information, solutions that might have already been solved either within or outside the organization are reinvented. For example, if within a specific research project a software method for the spiral-CT image segmentation has been developed which could be reused in another research project, it is normally re-implemented as its availability remains unknown. In order to prevent this, information and knowledge on existing solutions have to be systematically stored and easily accessible in order to enable their reuse.

Repeated mistakes: Another problem is that often mistakes are repeated due to the inability of organizations to identify or transfer lessons learned, e.g., using Fourier routines for problems where only Wavelet techniques have the power to perform effective texture analysis. In order to prevent this, tacit knowledge on solution strategies applied in the past to solve problems has to be captured explicitly and be accessible when required in order to provide expertise that already exist within the organization.

Resources: Academic environments have limited financial and human resources, which for example does not allow the establishment of additional infrastructural support or special training of new group members.

Constant evolution: The software R&D domain is characterized through continuous evolution of knowledge and technology advances. Therefore, relevant information and knowledge has to be continuously built and updated by constantly gathering new experiences each time a research project is planned and executed.

These goals and characteristics of Software R&D organizations in University environments show that comprehensive KM-based support is required.

3 Which Activities Need Knowledge-Based Support?

The principal activities of software R&D groups in an academic environment can be supported through knowledge and information:

- **Literature search and study** requires knowledge on where to find literature, to distinguish between relevant and irrelevant literature and on how to study literature effectively.
- **Writing research proposals** requires knowledge on the present state of art & practice and the specific research area and knowledge on how to formulate research proposals.
- **Meeting, presenting and discussing ideas** requires organizational support on scheduling meetings, knowledge on how to prepare and present research work and social skills wrt. discussions.
- **Development of theoretical models** requires knowledge on the state of art & practice and the application domain as well as knowledge on research methodologies, on similar research projects and on who, internal or external to the organization, has expertise in the specific area and on solutions which could be reused.
- **Development of prototypical software systems** requires knowledge on software engineering, the systematic development of prototypes as well as capabilities regarding the areas of interest and knowledge on who has expertise or worked on a similar problem and reusable solutions.
- **Writing and publishing scientific papers** requires knowledge on the state of art & practice, the application domain as well as writing skills.
- **Cooperating with research organizations**, industrial and application partners requires the integration of different background knowledge and the sharing of research results, (tele-) meetings and social communication skills.
- **Organization of research** or industrial projects requires knowledge on research programs, disciplines for the planning of projects and the execution of projects.
- **Attending conferences** requires information on upcoming conferences related to the research area of interest and funding possibilities.
- **Organization of events** requires knowledge on planning and execution procedures of events, funding and sponsoring possibilities, etc.

- **Teaching of knowledge** wrt. the research area of interest could be supported through overviews on information sources, tutorials, etc.

These requirements show that it would be beneficial to connect the individual researchers of the organization in a way that they could easily communicate and share any of the above listed type of information or knowledge.

4 Applying KM to a Software R&D organization in an academic research environment

As pointed out in Section 2, Software R&D organizations in university environments reflect learning on the organizational and individual level. Organizational learning is defined as a skilled process in which both explicit and tacit knowledge is created, acquired, and transferred wrt. the goals of the organization. During individual learning, the knowledge of one single person expands, in a process, in which experiences are transformed into knowledge, through observing phenomena, analysing them, developing models and theories about them and testing these theories an models in practice [14]. In order to effectively and efficiently support learning in university software R&D organizations, the processes necessary to continuously create and share knowledge across the organization have to be supported systematically. This is the goal of KM: to deliver the right information or knowledge at the right time, at the right place, in the right format, satisfying the quality requirements at the lowest possible cost [22]. In order to operationalize KM in software R&D organizations in university environments, relevant know-how has to be continuously build up by gathering new explicit and tacit knowledge during the planning and execution of R&D activities. A logical and physical structure for the continuous build-up of know-how in a software organization is the Experience Factory (EF) approach [4]. The EF environment complements the project organization by enabling the continuous learning on software development from experiences from their software projects and corporate-wide communication of software know-how, thereby promoting the creation of organizational know-how and the establishment of core competencies of the organization.

The EF proposes an methodological framework for analysing and synthesizing all kinds of experiences, acting as a repository for those, and supplying these experiences to various software projects on demand. These organizational experiences cover all types of knowledge assets which can support the planning or execution of software projects. Besides internal knowledge sources including explicit knowledge, for example, expertise and lessons learned on software engineering technologies, documents on medical knowledge, deliverables (e.g. software code parcels) or process guides (e.g. on how to give presentations) as well as tacit knowledge in the heads of the researchers, the EF framework can be extended to external knowledge sources, such as the Internet, and in our specific context to knowledge gained through university lectures (see Figure 2).

For successfully building knowledge management capabilities within an organization in practice, four key areas have been identified:

- **Knowledge content**: identification and modelling of the valuable knowledge available to be used or exploited.

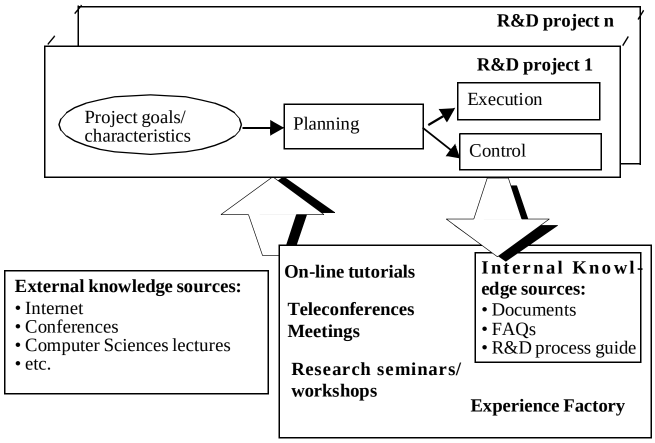

Figure 2 Experience Factory Organization in University Software R&D Groups

- **Organization**: structuring of the organization for knowledge exchange between human agents.
- **Processes**: definition of the processes to create, build, transform, organize, deploy and use knowledge effectively.
- **Technological support**: definition of tools supporting the creation and sharing of knowledge.

In the following sections, we describe each of these key areas outlining a KM solution customized to the characteristics of university software R&D organizations.

4.1 What Kind of Knowledge is Reusable?

A main goal of KM is to capitalize existing information and knowledge. This includes sources which are explicitly available in the organization, i.e., documents or tacit knowledge in form of personal experiences, which may be externalized, e.g., by inciting the members of the organization to write lessons learned on solution strategies applied or external knowledge.

The kind of knowledge relevant in software R&D activities can be determined based on the KM goals of the organization, which define the intended reuse through the object to be reused, the purpose for which the object is reused, the activity to be supported, the role, who reuses the know-how and the context in which reuse takes place. Examples of goals include:

- to **reuse documents** in order to support literature study from the viewpoint of a researcher in the research organization
- to **reuse lessons learned** in order to prevent the repetition of failures in the implementation from the viewpoint of a developer at the research organization
- to **retrieve similar FAQs** in order to create generalized guides for frequently reoccurring problems, e.g., on the understanding of image analysis techniques from the viewpoint of the knowledge engineer at the research organization Cyclops.

Based on the relevant knowledge in R&D activities (as described in Section 3) and the respective reuse goals, the following types of knowledge can be considered of interest:

Documents records (and files) constituting a personalized library including published literature, as well as, publications of members of the organization. The document records state referential information (such as title, author, year of publication) as well as comments and annotations (e.g., indicating the relevance of certain chapters). In addition, the record indicates the location where the document is available (e.g., university library) or, when available in electronic form, enables the upload and storage of the respective file. Such a library facilitates the access to relevant documents and helps to manage documents produced by the organization.

Project documentation including project proposals, plans and reports on research and industrial projects as well as students academic projects. The availability of those project documentations provides examples on plans facilitating the development of new ones and supports the execution of future projects.

On-line tutorials on topics relevant for the specific research environment. The availability of on-line tutorials enables researchers to learn on the specific topic independently and whenever necessary.

Frequently Asked Questions (FAQ) state a question and its answer provided by an expert. A collection of FAQs permits to automatically answer repetitively occurring questions. This facilitates the access to practical knowledge (e.g., on how to create a database connection to an application) and liberates experts to focus on more complex problems. A form to represent lessons learned in an accessible form is to describe them also in form of FAQs, by indicating a question which could be answered by the lesson and describing the lesson as answer to this question.

How-to-do recipes describing step-by-step how repetitively occurring tasks have to be done (e.g., how to connect a laptop to a X-ray computer). These recipes enable inexperienced researchers to perform these tasks without requiring the assistance of an expert.

WWW maps listing and commenting relevant Web sites. WWW maps guide the location of WWW sites of interest. They allow a more efficient access to relevant Internet sites that otherwise would be difficult and time-consuming to find.

Yellow pages capture human-resource capabilities and indicate experts wrt. specific research topics. The Yellow pages allow members to know whom to contact when they otherwise cannot find a satisfactory solution.

Starter´s kits summarizing and commenting information or knowledge relevant to a beginner wrt. a specific research area (e.g., introducing the programming language Smalltalk by referencing introductory texts, a tutorial and basic technical manuals, and indicating mailing lists and human experts). These pages allow newcomers to effectively and efficiently discover the respective research area.

News messages are any notes or comments of relevance to the research organization (e.g., to notify the availability of a new software version). The systematic management of these messages enables the communication of news and keeps the members informed.

Software parcels packaging and commenting software code or executables which have been developed within or are available outside the organization. Reusing these parcels prevents the re-implementation of readily available software and, thus, can reduces the development effort and improve the quality of the software.

R&D process guides are process models or guides on software R&D activities describing the activities to be performed (e.g., prototype development), model the outputs to be produced (e.g., on the structure and content of a master thesis) and describes how to do the activities. (e.g., on how to write a paper).

Conference calendar is a calendar indicating and commenting conferences of relevance to the research organization. The explicit availability of conference information facilitates the planning of conference participations and keeps the members informed.

The information and knowledge represented in the CM cover all relevant research areas wrt. the specific R&D organization, such as application domain (e.g., radiology, human anatomy), computer science and software engineering (such as image interpretation algorithms, software process model) as well as research methodologies (e.g., on literature revision).

4.2 KM Process in R&D Environments

The KM process can be basically divided into three phases [7]: knowledge generation, knowledge codification and coordination, and knowledge transfer.

Knowledge Generation aims at the generation of knowledge from various sources. Knowledge can be generated through knowledge acquisition from external sources such as conferences or through university courses related to the specific research areas. Inside the organization, knowledge can be created by bringing together people with different perspectives via research seminars, meetings or teleconferences.

Knowledge Codification and Coordination aims at putting organizational knowledge into a form that makes it accessible to those who need it. Its purpose is to locate and capture and turn organizational knowledge relevant wrt. the knowledge management tasks explicit and to organize and represent it in a form which allows its access, delivery and manipulation. Knowledge capture requires systematic procedures for acquiring, organising and structuring organisational knowledge. Knowledge from external sources can, for example, be acquired each time a group member finds a WWW site or a scientific paper of interest on the Internet. Knowledge acquisition from internal sources can include the exploitation of explicit knowledge already available in the organization (e.g., technical reports or project documentation) or to externalize tacit knowledge and to capture it in some kind of knowledge representation. For example, tacit knowledge can be externalized by writing down lessons learned or by explicitly describing research projects and the expertise of group members. For frequently re-occurring problems, how-to-do recipes can be developed as well as process guides. Knowledge can also be captured by documenting human-based knowledge management. For example, each time a question is answered by a human expert a possibility to acquire new problem-solving knowledge arises.

The acquired knowledge has to be represented and stored to make it accessible and useable to people in an organisation by explicitly specifying knowledge objects and their relationships of relevance. It may takes many forms in different situations, with variations in emphasis and formalisms, such as knowledge maps, hypertext, cases, etc.

Knowledge Transfer aims at developing strategies to encourage and enable the exchange of knowledge within the organization. Knowledge transfer can be done through person-to-person communications, e.g., through workshops, or through a distributed environment which allows to access relevant information and knowledge across the organization.

Executing this knowledge process in a cyclical way results in rapidly growing, well-evaluated, up-to-date, and demand-oriented learning contents of corporate learning systems [1],[5]. Besides these basic phases of KM, in order to cover a continuous knowledge life cycle, maintenance and evolution of knowledge and knowledge management processes is explicitly needed. The maintenance and evolution of the KM processes aims at continuously improve the provided support based on feedback from its application in practice (e.g., regarding the researchers' satisfaction).

4.3 Organization of KM in R&D Environments

Effective knowledge management requires a solution integrating management and technological infrastructure.

4.3.1 Technical infrastructure

To operationalize the Experience Factory in practice, KM requires a technical infrastructure, denoted Corporate Memory Management System (CMMS) [11]. A CMMS includes Corporate Memories (CM) which are the central repository of all the information and knowledge relevant to the KM task in an organization as well as tools to manage this knowledge-base. The tools should support the access to the right information or knowledge, the continuous acquisition of new experiences and their integration and storage as well as the continuous adaptation and maintenance of the CMMS to the specific environment..

Today, a wide range of Information Technologies are being used to implement CMMSs in general [2], as well as for the software domain [19]. These technologies include e-mail and group support systems, databases and data warehouses, browsers and search engines, intranets and internets, knowledge base systems and intelligent agents. Although, the approaches are numerous, none of these technologies itself offers a comprehensive support. In order to develop a CMMS that effectively contributes to KM in software R&D organizations, requires the integration of various approaches. For example, RetrievalWare [20] integrates natural language concepts and keyword searching and profiling. IKnow [12], includes natural language querying, mechanisms for the acquisition of new information and supports the continuous improvement of the retrieval performance. Other systems (e.g., Answer Garden [3]) integrate human experts into the KM process by forwarding questions which could not been satisfactory answered by the CMMS.

Based on our experiences, we propose a hybrid approach for an effective technical infrastructure to KM integrating techniques from various areas, such as Case-Based Reasoning, Information Retrieval and Information Filtering [9]. Our approach focuses on the support of the user through an intelligent assistant system, acquiring, providing, maintaining and distributing relevant information and knowledge (see Figure 3).

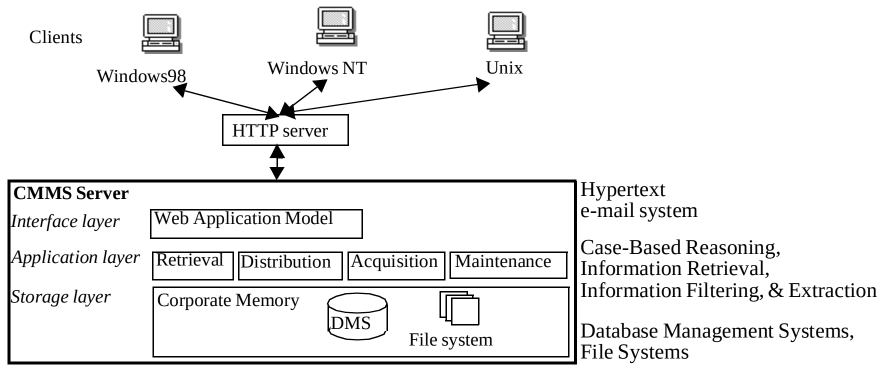

Figure 3 CMMS Architecture

Various types of reusable information or knowledge (as described in Section 4.1), denoted as CM assets, are stored in the CM. The CM assets are indexed wrt. their specific content as a basis for effective and efficient retrieval.

Besides the CM assets, general domain knowledge is represented in the CM defining terminology and basic concepts of the specific research areas (e.g., the programming language Smalltalk or human brain anatomy). It includes:

- **Ontology of relevant knowledge areas**: indicating relevant areas and their relationships (e.g., Smalltalk (Smalltalk database connection, Visualwave, distributed Smalltalk), Telemedicine (PACS, Teleradiology, DICOM), etc.) enabling a hierarchical classification of the CM assets.
- **Vocabularies**: indicating indicative expressions of the research areas for the indexation of CM assets. For example, the vocabulary on the programming language Smalltalk includes, the terms "class", "SortedCollection", etc.
- **Thesauri**: indicating associative or hierarchical relations between terms in the given domain, e.g., in the context of the programming language Smalltalk the terms "Collection" and "SortedCollection" are considered as similar.
- **Bilingual Dictionaries**: indicating the translation of domain-specific terms. In our specific application, we focus on Portuguese-English dictionaries. Here, example terms are "class -> classe" or "OrderedCollection -> coleção ordenada".

In addition, general vocabularies on the Portuguese and English language are represented as a basis for spelling correction of natural language queries.

The CMMS provides manifold support enabling the access to various types of information, knowledge or human experts for various purposes (e.g., facilitate a research on the state-of-art or guide the solution of a programming problem), from different viewpoints (e.g., developer, medical researcher). Search mechanisms (such as e.g., keyword search, similarity-based search) enable the effective and efficient retrieval of useful assets or to guide the access to human experts. Based on the user´s query, the CM is searched and the most relevant CM asset(s) are returned. If the CMMS does not provide a satisfactory retrieval result, the user can automatically direct her/his query to a human expert of the respective research area via e-mail. Once the answer is available, it is automatically send to the user. By composing the user´s query and the answer provided by the expert, a new CM asset is created and integrated into the CM (see Figure 4).

The search mechanisms have to allow the formulation of queries in natural language (e.g., in Portuguese and English in case of the Cyclops Project) and enable multilingual retrieval (e.g., searching also for English assets for queries formulated in Portuguese). In addition, the CMMS also assist researchers in finding information and knowledge that satisfies long-term goals (e.g. being aware of new publications in the area of radiological 3D reconstructions) through the pro-active distribution of knowledge wrt. to an user´s specific interests (e.g., informing about new published papers wrt. her/his research area).

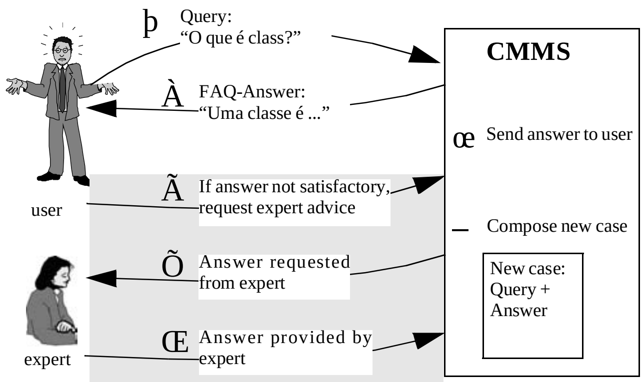

Figure 4 Interactive retrieval process

In the context of a software R&D organization, relevant information and knowledge is not completely available when creating the CM. Therefore, KM in software R&D organizations requires support for the continuous evolution of the knowledge base and has to be able to deal with incompleteness and inconsistency. This includes the continuous acquisition of new assets as integrated part of the R&D activities (i.e., reporting a FAQ on an recently occurred programming problem) and the indexing and integration of the new acquired assets into the existing CM. As the software domain is characterized through continuous changes and advances, the maintenance of the CMMS becomes especially important. This includes support for the maintenance of the knowledge assets, the general domain knowledge, as well as the improvement of the access mechanisms.

4.3.2 Management Infrastructure

The actual establishment of KM in Software R&D organizations needs to be embedded into a management infrastructure that provides funding and strategies to consolidate reuse and learning. Following the EF approach, the organization is divided into project organization and EF. The project organization includes several R&D projects. Within these projects, information or knowledge from external knowledge sources or the EF is reused and new gathered experiences are collected. In contrast, the EF aims at enabling the reuse of information and knowledge. Therefore, it focuses on supporting the generation, codification, coordination, sharing and maintenance of the knowledge and provides a technical infrastructure for the performance of these tasks. In accordance with the EF approach, every member of the research group has the responsibility to generate, collect and use available information and knowledge relevant to the research project. For example, a member of the Cyclops project who discovers how to establish a database connection within a Smalltalk image has the responsibility to write a lesson learned describing the required steps or a student working on her/his master thesis has the responsibility to include all literature references studied into the CMMS. In addition, a knowledge engineer, responsible for the creation and maintenance of the EF is required, who analyses, revises, evaluates and improves the EF. S/he is also responsible for initiating the development of more generic knowledge based on the observation of frequently re-occurring problems. The knowledge engineer also supports the researchers on request in retrieving and acquiring experiences.

A major problem regarding the introduction of KM in practice is the motivation of the members of the organization to explicitly document and share their knowledge. In order to be successful, KM activities related to the documentation of information and knowledge have to become part of the regular R&D tasks of the organization requesting from all members to share their experiences. This includes for example, the continuous documentation of studied literature or discovered web sites of interest. Students have to present what did work and what did not in their research, diagnosing the problems they encountered and the solution strategies they applied in research seminars, as well as by documenting project reports. In addition, regarding each's specific area of expertise, they are responsible to answer questions from their colleagues, which are captured as FAQ documents.

5 Outlook and Future Research

For continuous learning in a software R&D organization, systematic knowledge management customized to the specific characteristics and needs is important. In this paper, we describe the requirements for appropriate KM support in software R&D environments based on our experiences in the Cyclops research project. We describe the key areas of KM regarding the required content, processes, organization and technical infrastructure. Currently, we are establishing an EF in the Cyclops project. This includes the organization of weekly research seminars and semestral workshops, regular research topic-specific discussions as well as the implementation of a technical infrastructure. The current implementation of the CMMS focuses on the retrieval (search by navigation and attribute search) and the collection of documents and WWW sites. In addition, mechanisms for the handling and retrieval of FAQs via content-based searching on Portuguese natural language queries have been implemented [10] as well as techniques for the maintenance of the general domain knowledge (including Portuguese domain vocabulary, dictionary and thesaurus on Smalltalk issues). First evaluation results after one month of application in the Cyclops project (limited to the research group at the Federal University of Santa Catarina with 21 researchers) show the acceptance of the CMMS, with a total number of 330 accesses and 189 CM assets registrations. Based on further results of its performance and perceived usefulness, we intend to continue the implementation of the CMMS and to broaden the scope of research areas covered. In addition, we are also developing a R&D process guide for the Cyclops project based on the Quality Improvement Paradigm integrating incremental prototyping and extreme programming concepts. Future research also includes the development of intelligent knowledge-based techniques regarding information extraction, filtering and retrieval.

References

[1] K.-D. Althoff et al. Systematic Population, Utilization, and Maintenance of a Repository for Comprehensive Reuse. In G. Ruhe, F. Bomarius (eds.) Learning Software Organizations. Springer Verlag, 2000.

[2] A. Abecker, S. Decker. Organizational Memory: Knowledge Acquisition, Integration, and Retrieval Issues. Proc. of the Workshop on Expertsystems, Germany, 1999.

[3] M.S. Ackerman, T. W. Malone. Answer Garden: A Tool for Growing Organizational Memory. Proc. of the Conference on Office Information Systems, Cambridge, MA, 1990.

[4] V. R. Basili, G. Caldiera, H. D. Rombach. Experience Factory. In John J. Marciniak, ed., Encyclopedia of Software Engineering, vol.1, pp. 528-532. John Wiley & Sons, 1994.

[5] A. Birk, F. Kröschel. A Knowledge Management Lifecycle for Experience Packages of Software Engineering Technologies. G. Ruhe, F. Bomarius (eds.) Learning Software Organizations. Springer Verlag, 2000.

[6] The Cyclops Project (www.inf.ufsc.br/cyclops)

[7] T. H. Davenport, L. Prusak. Working Knowledge - How Organizations Manage What They Know. Harvard Business School Press, 1998.

[8] P. F. Drucker. The Post-Capitalist Society. Butterworth-Heinemann Ltd., 1993.

[9] C. Gresse von Wangenheim, D. Lichtnow, A. von Wangenheim. A Hybrid Approach for Corporate Memory Management Systems in Software R&D Organizations. Proc. of the Int. Conference on Software Engineering and Knowledge Engineering, Argentina, 2001.

[10] C. Gresse von Wangenheim, A. Bortolon, A. von Wangenheim. A Hybrid Approach for the Management of FAQ Documents in Latin Languages. Proc. of the Int. Conference on Case-Based Reasoning, Canada, 2001.

[11] C. Gresse von Wangenheim, C. Tautz. Summary of the Workshop on Practical Case-Based Reasoning Strategies for Building and Maintaining Corporate Memories at the 3. Int. Conference on Case-Based Reasoning, Germany, 1999.

[12] IKnow (www.knowlix.com/products/iknow.htm).

[13] KnowledgeX (www-3.ibm.com/software/swprod).

[14] D.A. Kolb. Experiential Learning. Prentice Hall, 1984.

[15] J. Laffey, R. Machiraju, R. Chandhok. Integrated Support and Learning Systems for Augmenting Knowledge Workers. Proc. of the World Congress on Expert Systems, Orlando, FL, 1991.

[16] I. Nonaka, T. Takeuchi. The Knowledge-Creating Company. Oxford University Press, Cambridge, UK, 1995.

[17] D. O´Leary. Knowledge Management Systems: Converting and Connecting. IEEE Intelligent Systems, May/June 1998.

[18] C.K. Prahalad, G. Hamel. The Core Competence of the Corporation. Harvard Business Review, 68(3), May 1990.

[19] G. Ruhe, F. Bomarius (eds.) Learning Software Organizations, Springer Verlag, 2000.

[20] RetrievalWare.(www.excalib.com/products/rw/index.shtml).

[21] P.Senge. The Fifth Discipline. Currency Doubleday, 1990.

[22] K.M. Wiig, R. de Hoog, R. van der Spek. Supporting Knowledge Management: A Selection of Methods and Techniques. Expert Systems with Applications, 13(1), 1997.

[23] K. M. Wiig. Knowledge Management: Where did it come from and where will it go?, Expert Systems with Applications, 13(1), 1997.