TECHNISCHE UNIVERSITÄT MÜNCHEN
Fakultät für Informatik
Lehrstuhl für Wirtschaftsinformatik (I 17)
Univ.-Prof. Dr. Helmut Krcmar

Fostering Knowledge Exchange Based on Relevant Activities

Jörg Schmidl

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Florian Matthes

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Helmut Krcmar
2. Univ.-Prof. Dr. Johann Schlichter

Die Dissertation wurde am 22.11.2011 bei der Technischen Universität eingereicht und durch die Fakultät für Informatik am 26.04.2012 angenommen.

Zusammenfassung

Ziel: Im Rahmen der Dissertation soll ein Ansatz entwickelt werden, der die direkte Wissensvermittlung zwischen zwei oder mehr ggf. unbekannten und entfernt arbeitenden aber im selben organisatorischen Kontext tätigen Personen unterstützt. Der Ansatz fokussiert auf Prozesse die durch IT-Systeme unterstützt werden und berücksichtigt die Vor- und Nachteile bestehender Ansätze. Ebenso fließen organisatorische Rahmenbedingungen für Wissensmanagementansätze, Mechanismen zur Motivationssteigerung, zur Minderung der Informationsüberflutung und zur medienbruchfreien Interaktion in die Konstruktion des Ansatzes mit ein.

Methode: Die Arbeit basiert auf einem gestaltungsorientierten Ansatz. Ausgehend von den Unzulänglichkeiten bestehender Unterstützungsansätze für die Wissensvermittlung werden Anforderungen an eine mögliche Lösung herausgearbeitet. Die Gestaltung des zentralen Artefakts wird zudem von Theorien aus Psychologie und Soziologie sowie grundlegenden Erkenntnissen aus dem Umfeld des Wissensmanagement geleitet. Um die Tragfähigkeit des Ansatzes zu belegen, wird eine prototypische Instanziierung des Ansatzes erstellt. Der Ansatz und seine Instanziierung in Form eines Prototypen innerhalb eines SAP System werden mittels deskriptiver Evaluation, Architekturanalyse und in zwei mehrmonatigen Fallstudien hinsichtlich der gestellten Anforderungen an das Konzept und der Vorteilhaftigkeit evaluiert.

Resultate: Aus der Artefaktkonstruktion in Form von allgemeinen Ansatz und prototypischer Umsetzung lassen sich folgende Resultate ableiten. Die alltagstaugliche Umsetzbarkeit des Ansatzes in bestehenden – auch großen – IT-Infrastrukturen ist nachweislich möglich. Die Nutzer gaben nach einer mehrmonatigen Nutzung in einer Befragung an, die Unterstützung sinnvoll zu finden. Dabei zeigte sich insbesondere, dass die Möglichkeit, Wissensbedürfnisse direkt aus dem Kontext der Aktivitätsdurchführung zu erstellen für Wissenssuchende sehr relevant ist. Ebenso wurde es als Vorteil gesehen, die Adressaten nicht selber bestimmen zu müssen. Ein zentrales Detail stellt dabei das Ergebnis dar, dass kontextuelle Ähnlichkeit im Vergleich zum relativen Status als Experte, als wichtiger erachtet wird – also aufgabenbezogene Verbundenheit hoch relevant ist. Durch sie ist eine höhere Anzahl an Interaktionen zu erwarten und der Wissenstransfer gestaltet sich leichter. Da das Maß zur Bestimmung der Ähnlichkeit eines Wissenssuchenden zu einem Wissensträger vielfältig gestaltet sein kann, wurde im Rahmen der Arbeit zudem eine adäquate Werkzeugunterstützung geschaffen. Mittels eines Trainingsdatensatzes kann damit ein passendes Ähnlichkeitsmaß bestimmt werden.

Auswirkungen auf Praxis: Die Vermittlung von Wissenssuchenden und Wissensträgern ist im Wissensmanagement ein oft adressiertes Problemfeld, dass jedoch noch nicht zufriedenstellend gelöst wurde. Diese Arbeit stellt eine neue Herangehensweise dar, die als eines seiner Designkriterien die Minderung bekannter Herausforderungen enthält. Für Organisationen, die den Ansatz umsetzen, ergibt sich nach der Evaluation der Ergebnisse – insbesondere der Fallstudien – ein zu erwartender Mehrwert. Der hierfür genutzte Prototyp wurde in einem SAP System umgesetzt. Seine Tragfähigkeit ist im wirtschaftlichen Umfeld wegen der starken Verbreitung dieser Systeme besonders relevant.

Abstract

Purpose: In the context of this dissertation an approach shall be developed that supports the direct knowledge transfer between two or more individuals who possibly do not know each other, might be situated in dispersed locations but who operate in the same organizational context. The approach focuses on processes that are supported by IT systems and takes into account the advantages and disadvantages of current state-of-the art approaches. Additionally, organizational requirements for knowledge management approaches, mechanisms to increase motivation to participate, to limit information overload and to interact without change of media are considered during the design of the approach.

Methodology: The research effort relies on the methodological foundation of Design Science. Based on the inadequateness of current state-of-the art approaches for the mediation of knowledge transfer, requirements are derived that a suitable solution must meet. The construction of the central artifact is furthermore guided by theories from psychology and sociology as well as findings from the knowledge management discipline. To prove the adequateness of this thesis' approach it is prototypically instantiated. The approach and its instantiation in the form of a prototype within an SAP system are evaluated with the help of a descriptive evaluation, an architecture analysis and two case studies that lasted several months to find out to what extent the requirements could be met and how advantageous the approach is.

Results: From the artifact construction in the form of general approach and prototypical implementation one can conclude the results. It could be proven that the implementation of the approach is suited for daily use in existing IT infrastructures – also large ones. After being exposed to the approach for several months, the users indicated in a questionnaire that the offered support was very welcome. They especially indicated that the possibility to satisfy their need for knowledge from within their work context was very relevant. Also, they felt it an advantage that the respondent had not to be determined by them manually. A central detail can be found in the result, that contextual similarity was seen as more important in comparison to the relative status of expertise – hence task-centric relatedness is highly relevant. With it, a higher number of interactions can be expected and the knowledge transfer is facilitated. Since the measure to determine similarity of knowledge seeker and knowledge bearer may take various forms, in the context of this thesis an adequate supportive framework was developed. Using a training data set it allows to determine a suitable similarity measure.

Practical Implications: The transfer of knowledge between knowledge seeker and knowledge bearer is an often addressed field in knowledge management that however has not produced fully satisfactory solutions so far. This thesis provides a new approach that features as one of its design criteria the mitigation of known challenges. Organizations that incorporate the approach may expect a benefit in knowledge transfer considering the results of the evaluation – especially the results of the case studies. The prototype that was employed was implemented in an SAP system. Its advantages are of specific relevance in economic settings due to the high pervasiveness of these systems.

Table of Contents

Table of Figures

Table of Tables

Table of Acronyms

CKO .... *Chief Knowledge Officer*
CoP .... *Community of Practice*
CSF .... *Critical Success Factor*
CWAD .... *Common Workflow Audit Data*
ERP .... *Enterprise Resources Planning*
ERS .... *Expert Recommender Systems*
HCM .... *Human Capital Management*
IS *Information Systems*
ISR .... *Information Systems Research*
IT *Information Technology*
KM .... *Knowledge Management*
KMS .... *Knowledge Management System*
PAIS .... *Process-Aware Information Systems*
SCT .... *Social Capital Theory*
SME .... *Small and Medium Sized Enterprises*
TMS .... *Transactive Memory Systems*
WfMC .... *Workflow Management Coalition*
XML .... *Extensible Markup Language*
YPS .... *Yellow Page System*

1 Introduction

"Der Anfang ist die Hälfte des Ganzen"
Freely translated: "The begin is half of the whole"
Aristoteles (384-322), Greek philosopher

In ancient times, the knowledge of how to master fire and how to craft the first simple weapons to hunt animals posed significant advantages for our ancestors to withstand the perils of their environment and has shifted the evolutionary selection process to our favor. Also, later in the history of mankind the possession of advantageous knowledge made some societies more successful than others. Often this related to knowledge about how to build more efficient weapons or how to be better at tactical warfare, i.e. how to coordinate the single individuals most effectively. But also other, more peacefully applied knowledge could determine the faith of a society. Some societies like the Vikings had knowledge about how to navigate on the sea in a way to withstand the harsh conditions and were presumably among the first to touch ground in North America. Others, like the Egyptians acquired knowledge about the approximate timing of Nil floodings and arranged their seeding times accordingly to harvest larger amounts of crops. Similarly, on the level of organizations, having necessary knowledge has always been a distinctive factor. In the middle ages guilds formed, to foster the knowledge of a profession and the ones who were not part of it had a significant disadvantage.

Throughout history, knowledge has been a major factor for distinction between successful and less successful societies or organizations. Today, many claim that this dependence of success on knowledge is even more pronounced especially in the more developed countries of this world as they gradually transform into *knowledge societies* (Willke 1998). In these societies, the "classical" production factors land, labor, and capital decrease in their strategic importance for a company in favor of the "new" production factor *knowledge*. Of course, the possession of knowledge has always contributed to the success of an organization, so it is not a new production factor per se, but it gained in relative importance. Krcmar (2010) identifies four major reasons for this. First, the products on the market are increasingly complex, may be bundled with additional services, and are often tailored to the specific needs of a customer, which requires a larger degree of understanding and hence a need for knowledge. Second, to gain higher market share in competitive industries, the lifecycle times have reduced considerably which requires a high rate of innovation that in turn depends on having the necessary "raw material", i.e. knowledge. Third, globalization of markets but also of value chains creates a challenge for companies and the ones who master the distribution of knowledge across geographical borders are in advantage. Fourth, fluctuations in work force have become more normal than a few decades ago, which gives those companies a relative advantage that are able to conserve the knowledge of their workforce.

Along with this change in relative importance of knowledge for the success of a company there also is a change in the role the single individual plays. Nowadays, workforce increasingly consists of *knowledge workers* and hence suitable support of their work becomes more important (Davenport 2011, p. 1). According to Drucker (1988) knowledge workers are specialists in their profession who govern their work on their own adapting their performance to feedback from their environment. However, owing to the increasing diversification and radical advancements in knowledge, their specialization is naturally limited to certain areas – there are no universal geniuses like Leonard DaVinci anymore.

This new role of the individual, with much higher sovereignty and specialized knowledge at his hand, has initiated a new view on what companies are and should do. From the point of view of an economist, Grant argues that the very existence of a company can be explained by its ability to offer a suitable mechanism for integrating knowledgeable individuals (Grant 1996). Taking the view of a sociologist, Quinn states that a company's "[…] raison d'être becomes the systematic coordination of knowledge and intellect throughout its (often highly disaggregated) network to meet customer needs" (Quinn 1992, p. 72). And Holsapple and Whinston (1987), taking an information systems perspective, see organizations as "[…] joint human-computer knowledge processing systems […] as a society of knowledge workers who are interconnected by a computerized infrastructure". The common theme in these definitions is that establishing and fostering suitable connections between knowledge workers is of the essence for companies today.

1.1 Motivation and Problem Context

Being able to support knowledge work is obviously a competitive advantage for a company. So the question, how to adequately support knowledge worker becomes important. Davenport argues that knowledge workers are typically supported following either of two approaches: the *free-access approach* and the *structured-provisioning approach* (Davenport 2011, p. 2), both having their pros and cons. In the former one, the general goal is to give knowledge workers access to as many and as diverse sources of information and assume that they will handle and integrate the information autonomously. In the later one, the information delivered to the knowledge worker is governed to a larger degree by structured processes and systems. While structured delivery is well-suited for tasks that follow a routine, pursuing a free access model assumes that knowledge workers know what information they can use, how to manage it and how to find it. However, Davenport notes that "[…] workers may know how to use technology tools, they may not be skilled at searching for, using, or sharing the knowledge."

(Davenport 2011, p. 4) and hence if possible some structure should be imposed to guide the knowledge workers. This is where Davenport sees the greatest potential for productivity improvements, indicating that finding ways to sensibly create structure where only the free access prevailed would be beneficial (Davenport 2011, p. 6). The kinds of tasks that are best supported by a free-access approach amended with some structure are those that either are fairly routine but need a larger degree of cooperation among employees and those that a complex in their nature and hence non-routine but typically rely on experts' judgment. However, the question how to create a suitable amount of structure and which means to use for this purpose, remain challenges to be addressed.

In both of the cases for which Davenport (2011) sees most potential for improvement, the interaction between knowledge carriers is of the essence. This is consistent with the *Bitkom Forecast for Knowledge Management 2007 – 2011* (BITKOM 2007) that, among other research directions, predicted an increased focus on approaches to better support direct, personal exchange of knowledge. Owing to the specialization and the resulting diversification of knowledge in companies, without additional support facilities, finding suitable experts is a challenge which is further fueled by geographical distance that also limits the ability to know who to address. Often this support is provided using IT solutions that are capable of reducing the limitations imposed by geographical distance and lack of awareness of other individual's knowledge. Therefore, approaches to foster knowledge exchange typically have an IT component. Also they typically focus on situations where people do not know each other and identification of a *suitable* interaction partner is challenging.

A first step when supporting knowledge transfer therefore lies in identifying the right interaction partners. When routine tasks require cooperation, indentifying the right persons to cooperate is necessary and, when relying on the support of an expert, finding the right one is also an important step. Some prevailing approaches try to facilitate the search for the individuals that possess the right knowledge for a specific situation. For example, directories containing the knowledge of individuals are setup that are then used similar to a telephone book. Another example can be found in knowledge communities that are established in a deliberate effort to create places where experts may exchange new thoughts or may help the less experienced members of this community.

While offering some structure and otherwise following the liberal *free-access approach,* those approaches often suffer from an inherent problem that comes with growth: the structure that is imposed is not selective enough to help the knowledge worker in the selection process of the right expert when in need of one. Reverting to the knowledge community example and looking at a specific one, this problem becomes apparent. The SAP Community Network (SCN), one of the largest knowledge communities world-wide, has far more than one million members, an average forum within the SCN can have 400 posts per month and there are more than 250 forums – a sign of a vividly used community but also one of information overload that renders finding the right expert a challenge. Many companies have internal communities that flourish similarly and hence face a similar challenge. Consequently some authors, e.g. Qureshi, Briggs et al. (2006), see an ongoing challenge to harness the power of knowledge networks mainly due to the fact that limited human attention puts a natural limit to the ability to collaborate. Ye, Nakakoji et al. (2008) also stress, that an employee's attention needs to be handled with care, finding a good balance between information overload and spreading sufficient amounts of relevant information within a network. In essence, employees want to stay in the information loop, want to be able to access the knowledge of others, but want these capabilities to be context-dependant to limit non-value adding attention loss. The prevailing approaches have some serious limitations to this respect and are incapable of providing advanced filtering mechanisms.

The effort that any knowledge management (KM) approach creates for the affected individuals should be as low as possible but in any case the possible benefits must outweigh the effort. This is a serious challenge for many current approaches. For example, keeping knowledge directories up-to-date typically inflicts extra burden that limits its acceptance, but the single individual is often not motivated to keep his directory entry up-to-date as there is no obvious benefit for him. Automation of tasks that are tedious for the individual but necessary for the approach are one way to relieve the knowledge worker and increase acceptance.

Another challenge that current approaches face, is their disconnectedness with operational tasks. Many tasks that a knowledge worker performs are either directly supported by IT, e.g. working on workflow items delivered to the individual from a workflow management system, or they are indirectly supported, e.g. when projecting future sales, adequate tools like an Enterprise Resources Planning (ERP) system's functionality is employed. Hence, for most tasks that a knowledge worker performs some IT-supported task is at least related to the original task. This allows for suitable integration of KM efforts, but current approaches often make little use of this. For example, directories that contain individuals along with their expertise are unaware of the current context of the individual out of which the need to use them arose. Similarly, knowledge communities are additional entities that are often not integrated into the operational tasks in which the demand to use them originates. On the other hand, embedding knowledge activities into the work context is an important requirement (Qureshi et al. 2006). It reduces the problem of having to change (communication) media and hence reduces the barriers for using any system and a system that helps finding the right persons in particular. Integration from an IT point of view is only possible in IT-supported operational tasks. As argued above, most knowledge-intensive tasks lead to the execution of IT-supported tasks and hence those tasks reflect the current work context of the knowledge worker suitably well.

In sum, there are important challenges to adequately support knowledge transfer. Imposing some structure for knowledge workers is beneficial but it is unclear how much structure and how to determine the right means. Also, having insufficient ability to keep up the awareness about all possible experts and having only limited, scarce amounts of attention, it is unclear who is a suitable interaction partner and how to best find one while keeping the effort for the individual low. Finally, a challenge lies in integrating supportive measures into operational tasks. The current approaches to support knowledge transfer fail to tackle at least some of the challenges. The approach developed in this thesis will address all those challenges and will also identify and cope with additional challenges that current approaches face. For this purpose, it utilizes the context of a knowledge worker as described by his currently exercised tasks in IT systems. The approach hence covers any task that is at least partially supported by an IT system.

1.2 Goals of this Thesis

Considering the observations of the previous section, it is this thesis' goal to describe an approach that fosters person-to-person knowledge exchange and that addresses the challenges that current approaches still face. More specifically, the following general design requirements are addressed by this thesis approach:

1. The challenges of *current approaches* are removed or their impact is decreased.
2. The scarce resource *attention* of each individual is treated with care.
3. The *search effort* for determining suitable interaction partners is decreased.
4. The *maintenance effort* of the approach is small.
5. The support mechanisms are *integrated* into the knowledge worker's operational tasks.

Consequently, a first goal lies in finding those aspects of current KM approaches that limit those approaches' support capabilities and that might be improved by a new approach to the problem domain. The main goal then, is to design an approach that fulfills the design requirements elicited before. Consequently, the KM approach developed in this thesis aims at connecting the *right* persons. This transfers the problem to finding a suitable mechanism to determine appropriateness of individuals in a context-dependant way. The overlap of relevant activities that the individuals have executed in the past, will be utilized for determining the appropriateness of an interaction partner. The assumption is that the knowledge seeker is best supported by someone who has a) *sufficient experience* with the respective task and b) has been or is in a *similar context* of the task. The first part influences the ability to help while the second part influences the willingness to help, as feeling or "being" similar to another individual increases the likelihood of identifying with the other and also with the other's needs. Therefore, this thesis' approach determines those individuals as *right* that have some repetitive element in common with the requester and hence that are somewhat similar to the requester. Having a common ground is essential for knowledge transfer to happen, which is why similarity is a proper concept for finding suitable interaction partners.

However, when focusing on knowledge-intensive work and the tasks within it as units of analysis, the question arises what exactly is the overlap or similarity of individuals and their contextual setting. This is a blurry concept because work processes are typically ad-hoc and the order of tasks is determined by the knowledge worker rather than by strict rules. Therefore, the central element in the KM approach is the similarity of users' knowledge of tasks, which is reflected by their previous execution of certain tasks that are stored in Information Systems (IS). However, there is not "the one and only" correct notion of similarity. Rather the "correct" interpretation strongly depends on the work setting, the organization, and other factors. Therefore, the determination of a suitable similarity measure for each instantiation of this thesis' KM approach is a challenging task on its own, especially because the variety of possible algorithms and configurations is large. A goal of this thesis therefore also lies in finding ways to support the task of determining a suitable notion of similarity.

The knowledge demand of a person correlates to what he is doing. Therefore, the need to engage in knowledge exchange with peers is determined by the activities that are relevant to each user. As argued before, an increasing amount of work is reflected in or at least supported by IS, which is why the usage pattern of these everyday tools is a valuable source of information that can be used to infer the user's competencies and knowledge needs. The traces, which are left by each user's activities, can be used to compare the user's context in order to find people that are likely to benefit from mutual discussion and sharing of knowledge. It is this research's guiding assumption that by considering the collection of interactions that all users build up, it is possible to determine *relevant topics*, to determine which persons benefit from their peer's *expert knowledge* by focusing on the activities that are *currently important* for the respective user's according to their *current and previous tasks*.

This thesis therefore adds value by deriving and developing an approach that builds upon the overlap of IT-supported tasks to find suitable interaction partners for knowledge transfer. Additionally, it describes an instantiation of the approach and the results that could be obtained in two case studies. Finally, it describes the design of a supportive tool for the determination of a suitable measure to determine task overlap.

1.3 Research Questions

To reach the envisioned goals, this thesis is guided by four research questions that subsequently inform the progress of this research. First, currently employed approaches to support knowledge transfer are investigated to understand their challenges and benefits, which both inform the design of the approach that is developed in this thesis. Additionally, this thesis investigates which aspects need to be taken into account when designing a KM approach, which helps to elicit further requirements for the design of this thesis' KM approach. Then follows the approach for knowledge transfer based on the overlap of IT-supported tasks that addresses the previously determined challenges. Subsequently, the core aspect of this thesis' approach, the determination of the overlap of IT-supported tasks is discussed explicating how to appropriately operationalize this aspect using a framework. Part of the discussion will be the elicitation of requirements and the derivation of a suitable design for this particular framework. Finally, the evaluation of the approach and the obtained results are described.

Research Question 1 Which requirements have to be taken into account for the design of an IT-supported person-to-person Knowledge Management approach?

In the context of the first research question, contemporary approaches that support person-to-person knowledge transfer processes are investigated. Yellow page systems, expert recommendation systems and knowledge networks pursue goals most similar to this thesis and currently enjoy wide-spread adoption which is why those systems are investigated. In turn, the analysis focuses on how each of the three approaches addresses this thesis' objectives, which advantages it offers and which disadvantages it faces. The result of the analysis of each approach results in a set of requirements that a more suitable approach should address. Additionally, on a more general level, influences on the success of KM approaches are investigated. As the output of this thesis is a specific application of KM practices, this makes sense. To this end, first critical success factors for KM initiatives in general are determined before more specific success factors for IT-supported KM initiatives are addressed. The result of this research question is a set of requirements that influence the design of this thesis' approach for fostering knowledge transfer.

Research Question 2 What are the necessary elements of a Knowledge Management approach that builds upon the overlap of IT-supported tasks?

Building upon the requirements and additional factors that were derived in the course of the first research question, in the second research question this thesis central design artifact is described: the approach to foster knowledge transfer. The derivation of the approach's design is guided by the following steps. First different design options that one may choose from are discussed and the alternatives that are adopted are described in detail. There, also the use of *task overlap* as connecting element between individuals is motivated. Next, in line with the concept of *theory-driven design* (Card 1989; Briggs 2006) theoretical models that inform the design are discussed. Building upon these insights, a first overview of the layers of the approach is given before each layer's design elements are detailed. The result of this research question is this thesis' central design artifact: the approach for fostering knowledge transfer.

Research Question 3 ***How can the core aspect of this thesis' Knowledge Management approach - the determination of the overlap of IT-supported tasks - be supported?***

The third research question's focus lies on describing a supportive functionality for the core aspect of the approach: a framework for the support of task overlap determination. Task overlap may be defined in numerous ways. However, the reuse of appropriate algorithms, given sufficient possibility to parameterize, seems possible. Motivated by these two aspects, a framework that reduces the effort for finding a suitable measure is envisioned. The design of the framework is guided by the following steps. First, requirements for this generic tool support are elicited and described. Then, the envisioned framework's architecture is concluded and each of its modules' features is described. Featuring a generic data format and support for transformations from arbitrary source data, the framework is highly adaptable and extendable. The result of this research question is the description of the framework supporting the determination of the overlap of IT-supported tasks.

Research Question 4 ***What effects can be achieved using this thesis' Knowledge Management approach and which implications for further research and development can be deduced?***

The goal of the fourth and last research question is the evaluation of this thesis' approach for fostering knowledge transfer. The evaluation focuses on different success measures. First, a criteria-based evaluation describes the relative benefits of the approach in comparison to the contemporary approaches that were analyzed. Then the implementability of the approach is evaluated showcasing its instantiation in an IT landscape that comprises more than 100 IS and many thousands of potential users. Finally, using the previous evaluation results two case studies were conducted. The first one served as general proof of concept in terms of user acceptance. Additionally, it served the purpose of getting first user feedback and to find implementation-specific suggestions for improvement. The second case study then showcased the usefulness of the approach in a larger, distributed real life setting. The obtained results also indicated areas for further improvements in terms of the implementation of the approach and the approach itself.

1.4 Research Design

Properly conducted research needs to be inter-subjectively comprehensible, i.e. a third party must understand how the research results were obtained, must be able to judge the quality of the research process and the quality of the outcome of it. To enable the community of researchers working in the same discipline to comprehend someone else's research, each research discipline relies on "their" commonly accepted "ways of performing research".

In the German-speaking region's Wirtschaftsinformatik discipline, two very general ways of doing research are generally accepted. The first is performing empirical research – either qualitative or quantitative. The goal of researchers following this paradigm is to understand phenomena of the real world by observing them and making sense of them, e.g. by proposing cause-effect models. While this is the predominant way of conducting research in the Anglo-Saxon discipline Information Systems Research(ISR)[1], in the German Wirtschaftsinformatik discipline the second general way of doing research enjoyed and still enjoys equally pronounced support. Here, the creation of new artifacts for and within the existing environment to reduce or remove an identified problem is the paramount goal. In contrast to the previous paradigm that wants to explain what is observable, the *science of the artificial* (Simon 1996) also often referred to as *design science* "[…] attempts to create things that serve human purposes" (March, Smith 1995, p. 253). These "things" can be of different nature. March and Smith (1995) differentiate between *constructs* that are used to describe a phenomenon, *models* that are used to communicate connections between the constructs, *methods* that describe how to arrive at a desired state and *implementations* that describe a real, "physical" instantiation of the before mentioned.

While the goal of creating an innovative artifact to improve the situation for a class of challenges is straight forward, the way to reach this goal is less obvious. Often, due to the fact that entirely new problems or known problems in new environments are addressed with the artifact that is to be designed, there is no or only little theoretical basis for the design (Hevner et al. 2004, p. 76). The creative design of the artifact may be informed by some contingent theories (Briggs 2006). However, more often than not new theoretical models will follow the application and the analysis of the possible impact of the artifact. A better understanding of the core of the problem and the actual mechanisms that are behind the working may lead to those new theories (Nunamaker et al. 1991). Therefore, design science is essentially a search process for a suitable solution to an important problem and the result of it is an artifact that has utility (Hevner et al. 2004, p. 80).

[1] Information Systems Reseach is the rough equivalent of the German Wirtschaftsinformatik discipline in Anglo-Saxon countries.

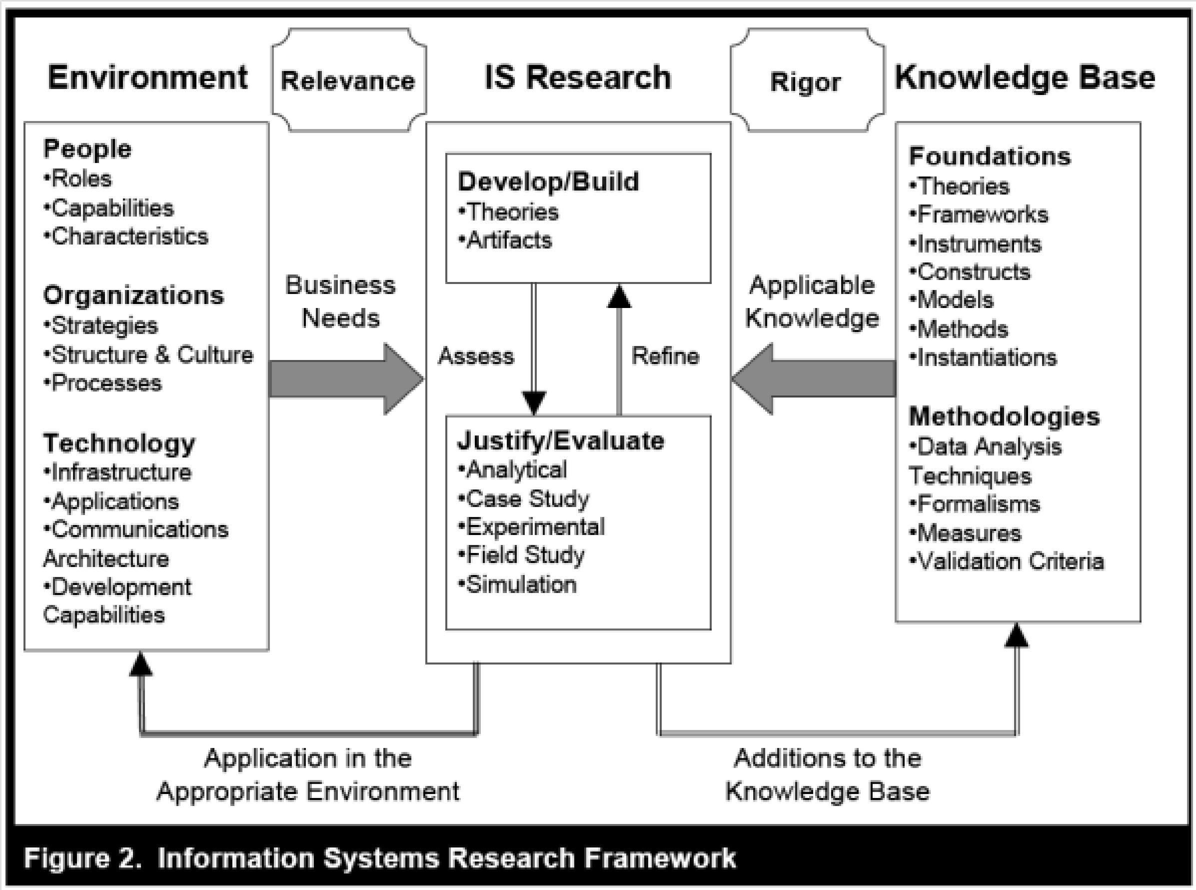


Figure 1-1 ***Hevner et al.'s Design Science Framework***
Source: (Hevner et al. 2004)

A summary of suggestions and a conceptual framework for this design process can be found in Hevner et al.'s discussion on how ISR should adopt design science (Hevner et al. 2004). Their understanding of the influencing elements while developing a suitable artifact is depicted in Figure 1-1. Design science inherently is an applied research paradigm and hence draws its purpose from actual needs that stem from the *environment*, which defines the problem space (Simon 1996). Hevner et al. (2004) adopt the very common structure in ISR that distinguishes between people, technology and the organization as the mutually influencing constituents of ISR (e.g. Silver et al. 1995, Krcmar 2010). The environment is also where the utility of the artifact will have to be proven which guarantees the relevance of the obtained results. As design science is a creative search process, at its core, there are iterations of artifact creation (upper rectangle in the middle) and evaluation of its utility (lower rectangle in the middle). The evaluation relies on suitable means, where suitability depends on the environment and the availability of previous results. During the design process, the *knowledge base* serves as a toolbox for the researcher. Accepted methodologies may inform the design process, existing theories may determine design choices and existing frameworks or instantiations may be incorporated into the new artifact. By building upon pre-existing knowledge from the ISR discipline and its related disciplines research rigor can be guaranteed (Hevner et al. 2004, p. 80). The artifact or experiences from its design and application may contribute to a better understanding and hence to e.g. better models about reality or more applicable frameworks that extend the knowledge base.

Taking Hevner et al.'s (2004) research framework, this thesis follows the design science paradigm. It addresses a current challenge in the business world, namely the still unsatisfying support for knowledge exchange in non-co-located work settings that poses a *Business Need* (right-facing arrow in Figure 1-1). To come to an innovative solution to this challenge, the thesis draws upon existing artifacts and the results that were obtained by their application to derive *Applicable Knowledge* (left-facing arrow in Figure 1-1) which is discussed in chapter 3. Also theoretical and empirical results that inform the design of a possible solution to the problem are included in chapter 4 and subchapter 5.2. An instantiation of the proposed approach is brought to *Application in the appropriate environment* which is discussed in chapter 7. The insights gained from the application and the summarizing conclusion of the thesis are *Additions to the knowledge base*.

Hevner et. al's (2004) research framework highlights the cornerstones of properly conducted design science research. While they also give recommendations on how to conduct proper design research they do not explicitly recommend a process to follow. This thesis therefore adopts a design process suggested by Pfeffers et al. (2006). They analyzed previous publications on design science to propose a six phase process model for design science research in which they distinguish the following phases (Pfeffers et al. 2006, pp. 89–92):

1. *Problem identification and motivation*: The goal of this phase is to define the research problem and motivate why a solution is beneficial. It should justify the pursuit of the solution and help to communicate the researcher's understanding of the problem. Therefore, in this phase the impact of a proper solution and the state of the art need to be analyzed and presented.
2. *Definition of objectives*: The goal of this phase is to rationally use the problem understanding for quantitatively or qualitatively inferring the objectives of the artifact that is designed. Therefore, in this phase knowledge about current solutions, their efficacy and their downsides needs to be worked out.
3. *Design and development*: The goal of this phase is to actually create the artifact in a way that fulfills the defined objectives by first designing the solution and then developing it. In this phase theories and methods from the knowledge base (see right hand side of Figure 1-1) are utilized to inform the design and development.
4. *Demonstration*: The goal of this phase is to show the utility of the newly created artifact by applying it to a suitable context and verify that it solves the expected problem.

Therefore, in this phase knowledge about how to apply the suggested solution to the selected setting is necessary.

5. *Evaluation*: The goal of this phase is to compare the achievable results of the suggested solution to the defined objectives. It is necessary to know about and select among suitable evaluation methods (see Hevner et al. 2004, p. 83 for a list of suggested methods) to show the efficiency and effectiveness with respect to the objectives. The evaluation may show the suitability or may suggest a re-iteration of the previous three steps to further improve the results.
6. *Communication*: The goal of this step is to inform relevant stakeholders, i.e. researchers in the same discipline and practioners that can benefit from the new artifact. The communication may comprise details about the problem domain and its significance, the suggested solution, its expected and possibly verified contribution and the rigor of the design process.

1.5 Structure of the Thesis

The support of knowledge exchange especially in industries and societies that increasingly create value by relying on knowledge as their most important asset, becomes ever more important. This thesis' goal is therefore to suggest a new and innovative approach that can support this process, i.e. to design a suitable new artifact. Consequently, following the design science paradigm, the thesis arranges its logical flow along Pfeffers et. al's (2006) process model to subsequently answer the research questions formulated in subchapter 1.3. The eight chapters and their relation to the four research questions are depicted in Figure 1-2. It also shows on its right side, the corresponding phases of design research that are addressed by each chapter.

Chapter 2 lays the foundation for this thesis. Fundamental terminology that is used throughout the remainder of the thesis is discussed. Various perspectives for formalizing the relevant terms are inspected and working definitions within the scope of this thesis are established. The discussion of the terms also serves the purpose to understand where this thesis fits in from the global perspective of the KM research discipline and where most scientific benefit may be generated. Chapter 2 also includes the introduction and discussion of a theoretical foundation for the relevant problem domain, the transfer of knowledge between individuals, is being looked at in this chapter. Implicitly this model serves as basis for the currently employed solutions for the problem domain and also serves as basis for this thesis' approach.

Chapter 3 analyzes the current state of the art of approaches for expert mediation. The goal is to explore the knowledge base for instantiations and methods and the environment for infrastructures and applications to understand the current situation. This in turn allows to adequately formulate the domain problems and defining the objectives. For this purpose, the KM approaches and KM tools that are used in general are sketched and the ones that are specifically suited to support the task of expert mediation are selected. Three widely used approaches fall into the latter category: Yellow Page systems, expert recommendation systems, and knowledge networks. Those are analyzed with respect to their principle functionality, advantages they provide and especially the problems they face which builds the basis for the problem understanding. The chapter concludes with ten general requirements for a solution that remedies the challenges identified before. They constitute the objectives of this research and form the first set requirements for the approach that is later designed

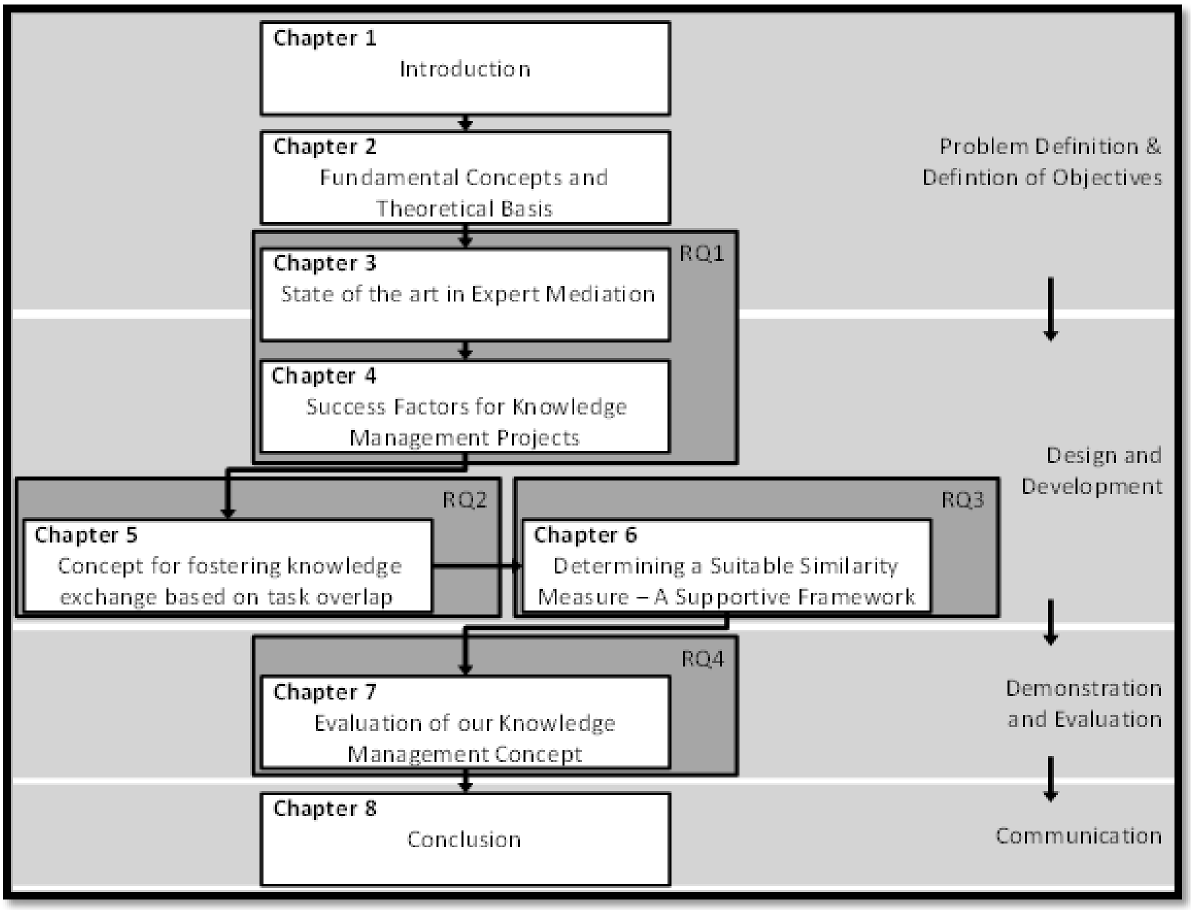


Figure 1-2 ***The relation of this thesis' chapters to the research questions and Pfeffers' design science research process***
Source: Own illustration

In chapter 4 factors that strongly influence the success of a KM project are investigated. Knowledge of these factors will inform the design of the approach in a way that increases the likelihood of attaining the goal of improved knowledge exchange. A literature analysis forms the basis to derive a common set of factors relevant to this thesis' goal. The resources that were used for the analysis build upon empirical research and hence exhibit the situation in the affected environment – companies that try to manage the asset *knowledge*. Also, existing models from the knowledge base that try to explain factors for success are utilized. From those sources a set of twelve general success factors for KM projects is derived. Additionally, a similar literature analysis was conducted that focused on success factors specifically for knowledge management systems (KMS) and their implementation that yields eleven such factors. The set of twelve general success factors and eleven success factors for KMS complements the list of requirements for the KM approach. Hence, together with chapter 3, this chapter answers the first research questions that asked for requirements and influencing factors that have to be taken into account for the design of a KM approach.

Having a proper understanding of the environment (chapter 2) and its challenges (chapter 3) and knowing general requirements for the design of a new approach (chapter 4), in chapter 5 the KM approach is derived. In a first step, the argumentation focuses on why a combination of two approaches – knowledge networks and expert recommendation systems – is a suitable approach to the addressed problem, before task overlap is described as the "glue" for the two approaches. Seeing the functionality of the two approaches through the lens of a combined approach yields additional design considerations that are elaborated in the course of this chapter. Further, adopting the idea of *theory-driven design* (Card 1989; Briggs 2006), two models that describe the exchange of knowledge from a social capital perspective and a socio-cognitive perspective, respectively are discussed. Taking those additional results into account an overview of the proposed approach is provided next. It consists of three layers: the *representation* layer, the *mediation* layer and the *data acquisition* layer that are explained in detail by describing their constituting elements. The representation layer fosters knowledge exchange by moderating the interactions among the users. While there are significant differences, for end users it resembles a knowledge network. In the mediation layer, the focus lies on using the context and expertise models of individuals to create context-dependent connections between them when knowledge exchange is sought for. The provided functionality strongly resembles the inner workings of expert recommender systems but deviates from it in important aspects. The data acquisition layer determines which data sources and which data entities are viable sources for the chosen approach. Also, the challenge of data extraction is addressed and a useful data representation along with a transformation application are described. The results of this chapter constitute the answer to the second research question that asked for the necessary elements of a KM approach that builds upon the overlap of IT-supported tasks.

Chapter 6 addresses the question, how appropriate tool support may be provided to facilitate the adoption of this thesis' approach. While most aspects of the approach need to be tailored to the setting, in which it is integrated, the task of determining task overlap, i.e. determining how similar two task contexts are, can be abstracted sufficiently to deliver a supportive framework. Requirements for a suitable framework capable of determining similarity of tasks contexts in various ways are therefore determined. Building upon the elicited requirements, the framework's architecture is proposed and its data model is described. Also, briefly the framework's usage of *model mining* is explained. It allows using non-formalized processes to deduce an implicit process model that can then be used to derive structurally similar task sequences. The result of this chapter constitutes the answer to research question three that asked how the determination of IT-supported task overlap can be supported.

Chapter 7 details the evaluation of this thesis' approach. First, possible ways for evaluating the approach are explained before the selection of methods and the reason for the choice are described. A descriptive evaluation followed by an architecture analysis form the first evaluation step. As a next phase of evaluation, the approach is instantiated for a specific use case: the facilitation of knowledge exchange in knowledge-intensive projects using SAP systems. Building upon the instantiation, a small-scale case study with some twenty participants and subsequently a large-scale case study with some 100 participants was employed to observe the adoption and the benefits that may be obtained. In relation to this thesis' goals, the descriptive evaluation shows that the approach does not suffer from the challenges of current approaches that were elicited in the course of this thesis. The two case studies verified that the instantiation of the approach imposes only small effort on the user and that the "consumption" of each individual's attention is reduced. Also, the integration of the implementation into the operational systems proved to be possible in an architecture analysis and was appreciated by the users as indicated by the results of the two case studies.

In chapter 8, the thesis concludes summarizing the results that could be obtained. Both groups of stakeholders – research and practice – are addressed. Also, this chapter discusses the limitations that were found for this thesis' results and consequently includes a call for further research to extend the obtained results and enhance the understanding which this thesis delivers.

2 Fundamental Concepts and Theoretical Basis

“Es ist nichts furchtbarer anzuschauen als grenzenlose Tätigkeit ohne Fundament.“
Freely translated: “There is nothing more abominable to see then eager work without foundation“
Johann Wolfgang Goethe (1749-1832), German Poet

Conducting research in a specific field requires a fundamental understanding of the field in beforehand. This way the correct terminology that has emerged over the course of time can be used to describe the phenomena that are investigated, general conditions that apply to the field of research can be taken into account and one may build upon the results that have been obtained in earlier endeavors and utilize the learning they provide. The contribution at hand concerns itself with fostering the exchange of knowledge and hence is part of the wider field of *Knowledge Management (KM)*. Consequently, the first part of this chapter will shed light on the interpretation of some fundamental terminology in relation to KM, before in the remainder, this chapter sketches the fundamentals of KM. The contribution at hand deals with a specific part of the KM, namely *Knowledge Logistics* (Lullies et al. 1993) and focuses on human-to-human interaction. This is why this chapter, also investigates a fundamental theoretical concept that originated in sociology – the *Transactive Memory System*. It addresses the actions that take place in a human-to-human interaction system with distributed pieces of global knowledge.

2.1 Knowledge

The exploration of the nature of *knowledge* and how it is obtained – called epistemology – has been the focus of philosophical debate at least since the heyday of the Greek philosophy (e.g. Platon n.d.). Still today, the question what knowledge *is* cannot be answered entirely, but different attempts to categorize and foster a better understanding have found their way into research. This subchapter tries to find a suitable interpretation of knowledge by first drawing upon the discussion from a philosophical point of view, before further narrowing down on the concept of knowledge by investigating different categorizations of knowledge. It then sheds light on the characteristic of implicitness of knowledge, the relation towards other terminology used in the field and categorizations that look at what “kinds” of knowledge exist.

2.1.1 Philosophical Perspective on Knowledge

A first definition dating back to Platon (n.d.) suggests that knowledge is *true*, *justified belief*. In this definition, belief refers to an opinion about reality. An opinion itself may be wrong (e.g. the opinion “I believe black and white are the same color” is always wrong) or right (e.g.

the opinion “I believe it will rain tomorrow” which may be true). The subset of opinions that are true constitute a first form of knowledge. However, a true belief also needs to be justified to be knowledge. Independent of the correctness of its meaning, the sentence “I believe it will rain tomorrow” is not justified and hence is not considered knowledge according to Platon’s definition. If the sentence is extended to read “It will rain tomorrow, because wind is pushing rain clouds towards this location” it is justified and would hence be considered knowledge. While, this definition sounds intuitively right, there is still defects in it from a philosophical point of view. Gettier (1963) showed that cases can be constructed where someone exhibits true, justified belief that, however, would not be considered knowledge and hence opposes Platon’s definition. In his constructed example, Gettier used a wrong assumption of his protagonist to derive a true result. There have been various attempts to change or extend the original definition to remedy this problem. One suggests to add, that the truth of the assumptions needs to be a fourth requirement (Armstrong 2008), another changes the definition to describe knowledge only if an opinion is believed and the justification for it ultimately forces truth (Kirkham 1984), which however limits the scope of what one may call knowledge to a very small set of items. Gödel (1931) formally proved that any consistent system for describing (a part of) the reality necessarily is incomplete and cannot describe everything that is true within the formal system. These results prompted some researchers to question the need of exactly defining what knowledge is altogether (Beckermann 2011).

2.1.2 Explicit vs. Implicit Knowledge

Dating back to Polanyi (1967) the distinction between *explicit* knowledge and *implicit* knowledge prevails to be the most common classification schema for knowledge. Implicit knowledge, also referred to as tacit knowledge, is person-specific, context-specific and not at all or only to a limited degree can it be communicated or put into a formal form. On the other hand explicit knowledge, also sometimes referred to as codified knowledge, may be cast into a systematic, transferrable “language” (Nonaka, Takeuchi 1995). Depending on the respective author’s understanding tacit knowledge may sometimes however also refer to not-yet-codified knowledge (Schütt 2003). While some influential researchers (e.g. Nonaka, Takeuchi 1995) interpreted the implicit-explicit dimension as a dichotomous classification schema arguing that transformation from one end to the other is possible, Tsoukas (1998) pointed out that both aspects are actually always present in every knowledge item – everything we know is always partly implicit and contingent to our previous experiences and partly explicable. This is why approaches that assume that storage of knowledge in a context-free way is always possible, which was the main theme in the beginning of IT-supported knowledge management, often failed (Schütt 2003). The more knowledge is context-dependant and/or hard to formulate the less it can be conveyed by efforts of formalization but rather it must be transferred by more direct forms of human interaction. This is consistent with the view of constructivism (Piaget 1996) – an epistemological view that assumes that there is no objective reality. There it is assumed that each individual builds up its own reality (and hence truth about the reality) by contrasting new impressions with previous experiences and making sense out of the result of this process. The own previous experiences and the mental model about the world is the intrinsically tacit dimension that only allows to absorb someone else's knowledge if the mental models are sufficiently in overlap – a process that takes place by mutual negotiation of interpretations (Lave, Wenger 1991).

2.1.3 Hierarchical Views on Knowledge

Another way to describe knowledge is by analyzing its relation to other terminology. Rehäuser and Krcmar (1996) describe knowledge as a network of information that constitutes some understanding of mechanisms, i.e. justification for a phenomenon. Information in turn consists of instances of data that are seen in a certain context. Data consists of a set of characters that follows syntactical rules. This hierarchy along with an example utilizing currency exchange rates is depicted in Figure 2-1. North et al. (2005) proposed a similar hierarchy which they called the *knowledge ladder*. It extends Rehäuser and Krcmar's hierarchy by introducing additional levels above *knowledge* with the goal to focus more on strategic aspects of knowledge management. Assuming motivation and application, knowledge may result in *actions*, if those actions constitute right choices they may lead to *competence* or *ability* that, in the highest level, may become *competitiveness* when it is unique to the individual or company.

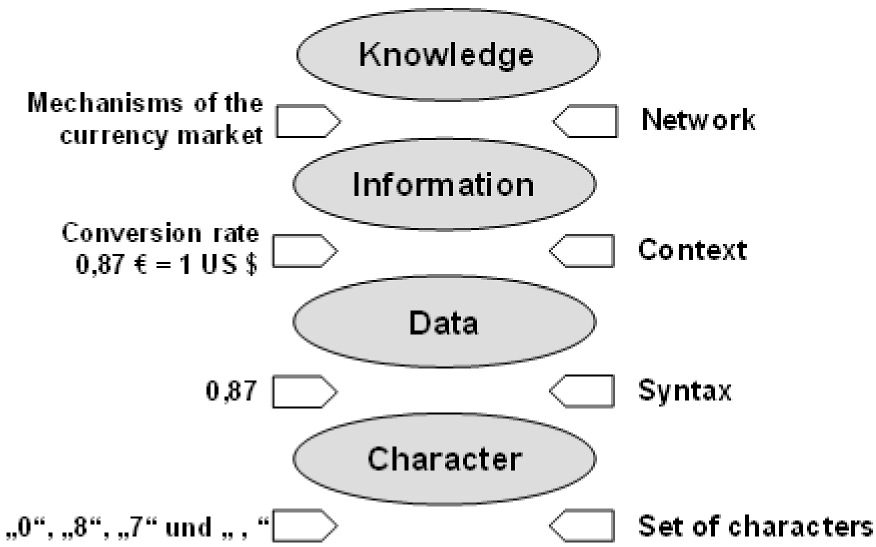


Figure 2-1 ***Illustration of the relation between character, data, information, and knowledge***
Source: adapted from (Rehäuser, Krcmar 1996)

Another definition that relates knowledge to information and data defines knowledge as „ […] the collective of skills and abilities that individuals use to solve problems. Knowledge is based on data and information, but in contrast to those is always bound to a person. ”(Gabler Wirtschaftslexikon 2011). According to this definition, explicit knowledge does actually not exist, as once knowledge is explicated it is information.

2.1.4 Knowledge Categories

The term knowledge may actually refer to different concepts that have in common that they enable an individual to deal with a task at hand. Therefore, some authors suggested to distinguish four kinds of knowledge (Lundvall, Johnson 1994; Jensen et al. 2007):

- Know-*what*
- Know-*why*
- Know-*who*
- Know-*how*

Know-what can also be termed factual knowledge. It refers to knowing numbers and facts and is very close to what other authors treat as information. *Know-why* refers to “[..] scientific knowledge of principles and laws of motion in nature, in the human mind and in society” (Lundvall, Johnson 1994, p. 27). *Know-who* constitutes knowledge about “[..] who knows what and can do what” (Lundvall, Johnson 1994, p. 27). Finally, *know-how* refers to skills and capabilities on a practical level, which is also sometimes called procedural knowledge. Jensen et al. (Jensen et al. 2007) note that the kind of knowledge also determines which “channels” may be appropriate to learn new knowledge. Know-what and know-why, often also referred to as *declarative knowledge*, are typically explicit knowledge and can be learned by reading books, attending lectures or accessing databases, while the other two are rather obtained by *experience learning* (Jensen et al. 2007, p. 682).

2.1.5 Interpretation Used in this Thesis

In this thesis, knowledge is assumed to be true, justified belief about the reality without further appreciation of the problems with this definition that is often less obvious in everyday use. While, as previously described, knowledge always has a tacit and an explicit part, this thesis will focus on such knowledge that exhibits more traits of its tacit dimension, i.e. it is not easily codified and highly context-dependant. Therefore, a later focus will lie on processes that deal with knowledge transfer by direct human interaction. Further, this thesis appreciates that knowledge is related other concepts such as information and data, a fact that will be exploited later when the suggestion for a way to connect knowledge bearers is derived. Finally, this thesis’ approach focuses less on the know-what and know-why categories of knowledge but tries to support the know-who category and, since it will exploit the interactions of knowledge workers with supportive IS, it focuses on the support of the transfer of know-how. A suitable fit of this thesis' interpretation can also be found in Davenport's definition of knowledge that suggests to see knowledge as

> "[…] fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the mind of knowers. In organizations, it often becomes embedded not only in documents or repositories but also in organizational routines, processes, practices, and norms." (Davenport, Prusak 2000a, p. 5)

2.2 Knowledge Management

This sub-chapter elaborates on the term K*nowledge Management (KM)* to come to a suitable understanding that builds the basis for this thesis' further research. To this end, it first looks at definitions common to the ISR discipline before sketching the development of KM over time to understand where new contributions to the field, like this thesis, will fit in. Finally, two models of knowledge management are addressed: the *SECI model* and the *KM task model* of Probst/Raub/Romhardt. Both are widely cited and discussed in the field and will be used in this thesis to classify the thesis' contribution in the larger area of KM research.

2.2.1 Definitions of KM

The discussion about what constitutes KM is similarly pronounced as the discussion about what knowledge is and as a consequence various attempts to describe the field can be found. This is especially the case as KM is truly a trans-disciplinary field of research with many research area's contributing to the matter. Maier (2007) describes different reference disciplines so do Götz and Schmid (2004) describing fundaments for knowledge management from research in sociology, marketing, human resources, market competition, creativity and innovation. Acknowledging this multi-disciplinary view on the complex matter of KM, next follows an investigation of attempts for defining KM from an ISR perspective.

From the ISR perspective, KM lies at the intersection of the three dimensions *organization*, *people* and *technology*, where each of the three can improve the situation with respect to knowledge use and flows. Using a broad definition to describe the phenomenon, Bellmann and Krcmar (2002) understand KM to include theories, methods and tools to enable a systematic use of the resource knowledge, hence KM subsumes every activity that deals with knowledge. Probst et al. emphasize the organizational benefit by defining knowledge management as „[..] an integrated concept of intervention that deals with the possibilities to create an organizational knowledge base" (Probst et al. 2006, p. 23; translated from German by the author). Krcmar adds details to the form of intervention when he argues that KM can be seen as "[..] the connection of innovative information technology and new organizational concepts to improve knowledge-intensive processes within a company." (Krcmar 2002, p. 26; translated from German by the author). Maier, in his interpretation of the term, emphasizes the strategic dimension of knowledge management by considering the proper selection, realization and evaluation of goal-oriented strategies as characterizing element for knowledge management (Maier 2007). In this thesis' context Krcmar's (2002, p. 26) definition is most suitable, as in the thesis, there will be innovative use of Information Technology (IT) that may create new organizational connections as a result with the overarching goal to improve the knowledge transfer in an organization.

2.2.2 Development of the KM Research Field

The focus of KM research has shifted over the years. While it appears natural to think of KM dealing with documents, people and social interaction (Oldigs-Kerber 2007, p. 65), the respective entities enjoyed different attention during the development of KM. However, it is worth noting that each new wave of KM did not replace the previous one. Merely the focus in research shifted and the available knowledge about KM was extended. A detailed account of the historical development of the KM discipline can be found in (Wiig 1997) or (Prusak 2001). In the following, the major "eras" are sketched to motivate why this thesis focuses on human-to-human knowledge transfer with emphasis on group interactions.

There were some early discussions about KM in the 1970ies, however, sociologists mostly drove these. The respective terms and ideas were not in the focus of management magazines until the early 1990ies, when all major consulting companies started talking about them (Schütt 2003, pp. 451–452). This was the emergence of the first wave of KM approaches that in retrospective is seen as too technocratic as the focus almost exclusively lay on supportive IT. Within this first wave, Knowledge Repositories, sometimes also called organizational memory systems, were considered as *the* solution for all problems. The general idea was to enable an organization to remember everything by storing knowledge and experiences in a database or similar data store. As the focus was storing electronic documents containing knowledge, this phase of KM is also sometimes referred to as document-centric (Oldigs-Kerber 2007, p. 65). However, those systems faced four major problems that lead to unfulfilled expectations towards them (Ackerman, Halverson 2004):

- The goal of having one single data store across an organization failed due to technical challenges of integration and the natural political rivalry within companies.
- The belief that every piece of knowledge could be sufficiently decontextualized or brought into a storable format at all turned out to be not the case.
- The assumption that people would share their knowledge spontaneously did not hold.
- The assumption that what one person enters into a repository would be automatically understandable for another person retrieving it turned out to be wrong.

In the second wave of KM, the persons responsible reverted from the erroneous belief that all knowledge could be codified. They rather realized that some knowledge could only be transferred from person to person (Schütt 2003, p. 452). The focus of KM research steadily shifted toward the question how to relate people and as one predominant part of this challenge, to suitable ways of locating expertise in an organization. In their simple form, those expertise locator systems are yellow page systems with a query front end, where one may rely on a list of fields of expertise and a link to those individuals that posses this expertise. In their more sophisticated form, expertise locator systems are often also termed expertise recommender systems. They automatically extract levels of expertise from their users by different means using various sources of implicit proof of expertise. Independent of the implementation details, these systems focused on managing knowledge by connecting people, which is why those KM systems are also referred to as *person-centric* (Oldigs-Kerber 2007, p. 65). However, the general class of expertise locator systems faces some challenges (Ackerman, Halverson 2004):

- It is difficult to keep the information about who knows what up-to-date.
- The proper classification of people, their skills, and expertise is still the goal of ongoing research.
- It is also difficult to design those systems to be dynamic enough to cater for change.

Oldigs-Kerber (2007, p. 67) adds, that a focus on person-centric KM leads to a relatively unmanaged network of knowledge flows between individuals.

The third wave of KM focuses on *(social) organizational structures* (Oldigs-Kerber 2007, p. 65). The implementation of an *electronic community*, also called *online community*, *virtual community*, *community of practice* or *electronic network of practice* took center stage in research and practice. Here, people with expertise meet other experts to exchange expertise, knowledge, and opinions and by this way are able to learn from each other. Ackermann (2004) states that the challenge with those communities lies in the assumption that people will join. Understanding what motivates individuals to join and emphasizing the possible benefits of such communities is therefore the key. An extreme form of virtual communities can be found in ad hoc-groups (Ackerman, Halverson 2004). These groupings can come together quickly, actively work on a problem, and then disband after they have finished. Analogies from organizational sciences point at crisis teams or tiger teams, but new technologies have offered the vision of supporting geographically distributed and extremely short-duration teams. However, Ackermann (2004) sees two major issues with this vision. First, it is difficult to assemble individuals in such a way that they form an effective team, and work together efficiently. This is even more difficult when it is a purely virtual team. Second, it is hard to find the *right* individuals with respect to their knowledge – the same challenge that expertise locator systems face (Ackerman, Halverson 2004).

This thesis can be considered part of the third wave of KM, as it focuses on knowledge transfer between individuals but takes into account the benefits of spreading the learning effects to a larger (social) group. While not the focus of this thesis, the support of ad-hoc groups as envisioned by Ackermann (2004), might be supported by this thesis' approach as well.

2.2.3 Task-Oriented Model of KM

One way to structure a discipline, is to inspect which processes are relevant to the discipline and to formalize them into a model. This also helps to categorize where new approaches, like the one developed in this thesis, will fit in. Suggestions for formalizations of KM processes along with their relationship among each other can be found in a number of publications, for example (Alavi, Leidner 2001), (Laudon, Laudon 2006) and (Probst et al. 2006, pp. 25–33), of which the last mentioned enjoys most attention in the German-speaking KM community. It is depicted in Figure 2-2. The model's eight process categories and their interpretation are as follows:

- **Knowledge Identification** is concerned with achieving *transparency and awareness* on suitable and task-oriented knowledge resources that are already available. Knowledge can be identified internally in the company or externally in its environment and may cover explicit or tacit knowledge.
- **Knowledge Acquisition** is concerned with the question which new knowledge should be *acquired externally*. The adoption of external knowledge can be done by e.g. licensing a patent, company acquisitions, or new hires.
- **Knowledge Development** is also concerned with extending the company's knowledge assets but by *internal development* rather than external procurement. At its heart, in this process category organizations analyze and assess how to deal with new ideas and how available expertise and the creativity of employees can best be utilized in the interest of the organization.
- K**nowledge Sharing/Distribution** is concerned with facilitating the *sharing of knowledge* within the organization and its *distribution* to the *right* person at the *right* time. Efficient, effective, and reliable knowledge distribution is the process of spreading knowledge that is already available to the company taking into account cultural, strategic, and technological requirements.
- **Knowledge Utilization** is concerned with making sure that knowledge that is available, is also *applied productively*, contributes to achieving organizational performance goals, and, thus, creates business value. For this purpose, barriers like the "Not-Invented-Here" syndrome have to be addressed.

- **Knowledge Retention** deals with the *storage and maintenance* of once acquired implicit and explicit knowledge. The goal is to guard the knowledge against loss due to e.g. organizational reorganizations or loss of knowledge carriers leaving the company.
- **Knowledge Assessment** is, in contrast to the previous processes and in line with the next process, a *strategic* process. It is concerned with *establishing goals* and providing direction to KM projects and focuses on determining which knowledge and competencies are to be developed and at what organizational levels.
- **Knowledge Goals** compromises methods to measure (the value contribution of) knowledge and its application and allows *assessing the success* of KM initiatives. In addition, it enables organizations to record lessons learned in KM initiatives and provides feedback on the organizational knowledge goals regarding their quality and degree of achievement.

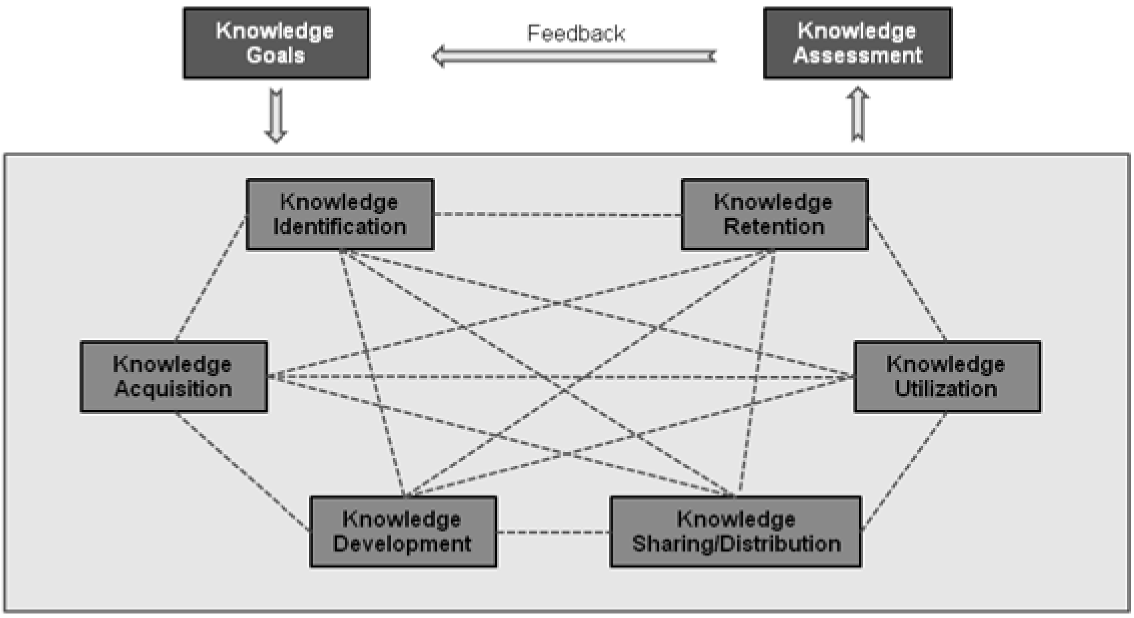


Figure 2-2 The Probst/Raub/Romhardt Model of KM Processes
Source: (Probst et al. 2006, p. 32)

This thesis will address the challenge to connect knowledge bearers and knowledge seekers to support the transfer of task-relevant knowledge. In this sense, it supports the KM process *Knowledge Identification*, as it helps to clarify who is knowledgeable concerning a specific task. Naturally, it also supports the KM process *Knowledge Sharing/Distribution* as the transfer of knowledge is the core of the approach. However, *Knowledge Retention* will also be a fundamental aspect addresses in this thesis' as the developed approach explicitly tries to proliferate the knowledge contained in one-on-one interactions. Finally, as a result of facilitated interaction, *Knowledge Development* is supported by this thesis' approach – sharing of knowledge among individuals leads to increased levels of knowledge within the social group, i.e. the organization.

2.2.4 The SECI-Model

Within the knowledge management discipline there is one model that, while sometimes discussed controversially, still belongs to the most-often cited and most influential models: the *SECI model*. Adapting Polanyi's (1967) thoughts about the tacit[2] dimension of knowledge to the challenge of organizational knowledge creation, Nonaka and Takeuchi (1995) proposed the SECI model that seeks to describe how organizations learn.

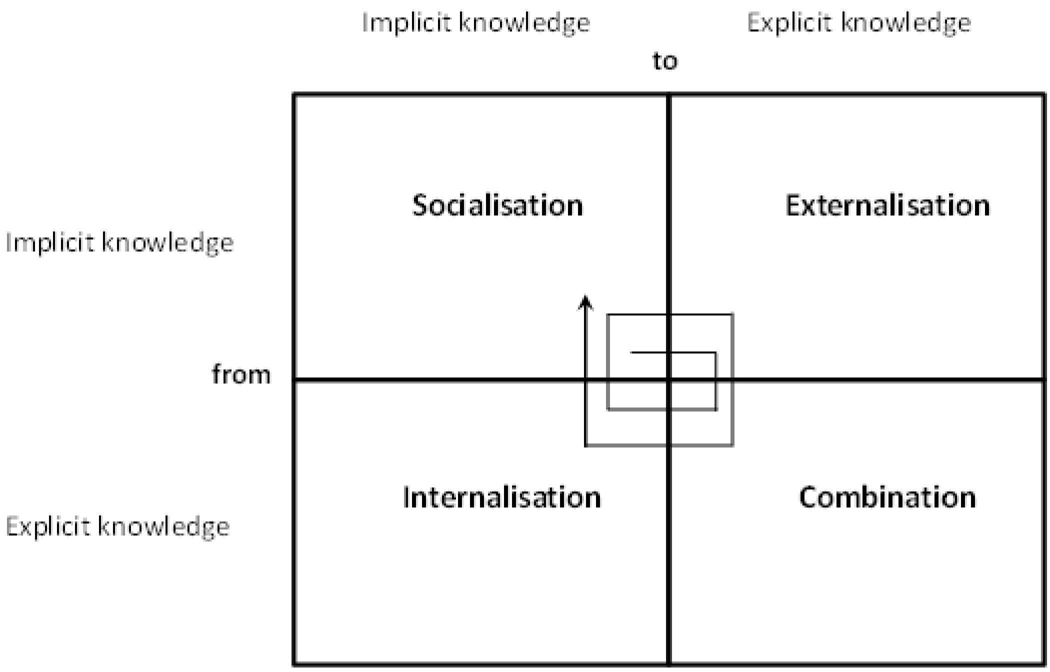


Figure 2-3 ***The SECI model's knowledge spiral***
Source: adopted from (Nonaka 1994, p. 62)

A core assumption of the model is that generation of new knowledge and hence learning effects for the company, primarily have their roots in the implicit knowledge of individuals that is then explicated. Nonaka therefore notes that "[…] the articulation of tacit perspectives in a kind of mobilization process is a key factor in the creation of new knowledge" (Nonaka 1994, p. 16). The other assumption is, that it is in principle possible to formalize tacit knowledge, which Nonaka acknowledges as challenging, when stating " […] to convert tacit knowledge into explicit knowledge means finding a way to express the inexpressible"

[2] Nonaka and Takeuchi speak about *implicit* knowledge and *tacit* knowledge interchangeably while Polanyi only refers to *tacit* knowledge .

(Nonaka 1994, p. 99). From these two assumptions and the distinction between implicit and explicit knowledge, Nonaka and Takeuchi distinguish four knowledge transformation processes that together form their *knowledge spiral*: **S**ocialization, **E**xternalization, **C**ombination, and **I**nternalization (see Figure 2-3).

Socialization refers to a process where knowledge is transferred from one individual to another in a non-formalized way by shared experiences or observational learning with apprenticeship being a typical application of this kind of knowledge transfer. The result is a shared knowledge of the group in which socialization took place – the group members have *shared mental models* that allow them to work together more productively. *Externalization* refers to a process where implicit knowledge is increasingly formalized using metaphors and analogies that eventually result in explicated models. The process is fueled by collective reflection about implicit knowledge. The newly created, explicated models can then, in the *Combination* process, be put together to amend each other or by sorting or categorizing the results of the previous phase which yields new knowledge in itself. The *Internalization* process refers to the acquisition of explicit knowledge by an individual, i.e. personal learning. This way, the knowledge base of an individual, subsequently the group that this individual belongs to, and later the whole organization is extended by iterating through these processes.

The SECI model has received major attention and formed the basis for many KM approaches and tools. Especially the first wave of KM (see section 2.2.2) tried to operationalize the idea that explicit knowledge may be transformed to implicit knowledge and vice versa. This, however, is also the major point of criticism for the SECI model. Many argue that explication is not possible for implicit knowledge in general and hence parts of implicit knowledge cannot be part of the suggested spiral, which renders the SECI model inappropriate at least in some cases (McAdam, McCreedy 1999; Wilson 2002; Schütt 2003). Nevertheless, this thesis' contribution can be categorized in relation to the SECI model to allow a better understanding of how it relates to previous approaches. It is this thesis' goal to support the knowledge transfer from individual to individual and hence the Socialization process is relevant. However , at the same time the approach developed in this thesis, tries to ensure that the learning that takes place during these interactions is made reusable to others at least to a certain extent and hence the Externalization process is relevant as well.

2.3 Transactive Memory Systems

This thesis addresses the challenge of adequately supporting knowledge transfer processes between individuals. Once individuals interact, they form a social fabric that may often be rather loose. A theoretical foundation for the interactions of a group may be found in Wegner's Transactive Memory Systems (Wegner et al. 1985; Wegner 1987). Therefore

Transactive Memory Systems (TMS) theory is described in the following before detailing how it influences this thesis.

2.3.1 Foundations of TMS

The TMS theory has its roots in the investigation of interactions in close relationships. Wegner expanded his results to groups trying to predict group and individual behavior based on his theory of information flows between the group members, which then is called a TMS. Formally it can be defined as "[..] a set of individual memory systems in combination with the communication that takes place between individuals" (Wegner et al. 1985).

An important element of the theory is the observation that information may be stored externally, i.e. outside of one's own mind, for example on a physical paper note but also "in" other people. Wegner notes that in this case there are additional prerequisites to reuse the information. "The successful retrieval of a memory item […] requires the prior encoding of at least two additional pieces of information – a retrieval cue or *label* for the item […], and a notion of the *location* of the item […]. This seems to be a general requirement for the use of external storage." (Wegner 1987, p. 188). Further reasoning about external memory, Wegner concludes that for retrieving memory items one either stores a label and the item to be retrieved or a label and the location of an external storage entity, i.e. a pointer to what is needed. In the second case, it is not necessary to know the memory item after all, as long as one knows how to get to it. If the external storage is not physical but is another individual one is dealing with TMS, where "[…] one person has access to information in another's memory by virtue of knowing that the other person is a location for an item with a certain label" (Wegner 1987, p. 189), so essentially *know-who* knowledge (see section 2.1.4).

2.3.2 Mechanisms in TMS

TMS theory builds upon an analogy to processes that take place in individual memories. Psychological models about learning and remembering of individuals, i.e. the individual memory (e.g. Zimbardo 1983, pp. 246–278), distinguish the three processes *encoding*, *storage* and *retrieval* of information items which Wegner directly extended onto the group level. *Transactive encoding* essentially is an explicit or implicit negotiation where in a group new information is to be stored, i.e. who is responsible to store it. This way one (or more) individuals acquire the new information itself – they are the experts on this matter – while the others obtain the information that the selected individual(s) know about this new information without knowing the information itself but only the label. *Storing* refers to the actual inclusion of the information into the group's collective memory, by storage in the mind of the respective group member. *Retrieval* refers to the process of getting to the information. If someone is an expert on the needed information item, it will be stored in his mind and he can answer right away. Otherwise, the *transactive retrieval* process starts. The information seeker may follow his “pointer”, i.e. the location information associated with the label to find people who possess the needed information. This process may also cascade, if the initial person that is addressed does not know but knows who to ask for the specific piece of information. Wegner stresses that there is a third way. During the search process in the group, the initial request for information may be *recoded* and *amended* by partial information of others in communicative acts. This may lead to finding the desired information, which would not have been possible with the knowledge of either individual alone but only through their interaction. In the same line of argumentation, Wegner notes that those integrative communication acts, i.e. two or more individuals combining the information items they have stored under the same label, may bring about new knowledge in the group (Wegner 1987, p. 197).

A TMS comes into existence, once individuals acquire knowledge about what others know, i.e. their domain of expertise and hence who is expert on what topic. During a TMS’s evolution, the expertness of individuals is often self-reinforcing. New information is directed to or implicitly assumed to be acquired autonomously by a dedicated expert. Consequently the expert is also seen as the one responsible for this area of expertise, which may become problematic for this individual as the workload may rise. Another aspect that determines who is seen as responsible in the group for harnessing specific information, especially if there is no formal assignment of this, is the *circumstantial knowledge responsibility* (Wegner 1987, p. 192). The one who first came across an area of expertise is, independent of the level of expertise, considered the “expert” for this topic. Also, the one who has last operated in this area of expertise may be considered the best suitable expert and may hence be seen as the person responsible for harnessing the respective information items.

2.3.3 Challenges of TMS

While the extension of accessible knowledge is an obvious benefit a TMS may offer to an individual, there are also challenges (Wegner 1987). As the knowledge about who knows what is the fundament, insufficient ability to find a path to the necessary information is a bottleneck. Therefore, in pursuit of information, it is essential to *find a suitable individual* to request information from, this person being the expert or someone who knows how to get to the expert, i.e. a dispatcher.

Another challenge lies in relying on the expertness of another individual which is an essential part of a TMS. It is, however, challenging to *determine the real expertness* of the other individuals and, as a result, of the group, which may in turn lead into overconfidence in the group’s knowledge.

Yet another challenge of a TMS lies in its vanishing. If the distributed storage of information falls away, the former members of the group are left with now *meaningless pointers* to knowledge items but this knowledge is no longer accessible to them.

A final challenge lies in the setup and design of TMS. Wegner (1987) distinguishes two extremes: *differentiated TMS*, in which each individual is an expert on its matter of expertise and there is virtually no overlap in expertise and *integrated TMS* in which the group has a substantial overlap in experts and many individuals know the same things. It is a matter of proper design in relation to the field of application to find a good balance between those two extremes.

2.3.4 Implications for this Thesis

The theory of TMS models *know-who* knowledge and how it influences the behavior of individuals within a group. Wegner's (1985) definition of a TMS does not address the group cohesion and hence any set of individuals that interact in the described way form one. However, taking into account the theory's roots in analyzing intimate relationships, the unit of analysis of TMS theory tends to be groups that are tied together by a "binding element" that for example motivates an individual to assume responsibility for a certain area of expertise for the group. In this thesis, the focus lies on supporting knowledge transfer processes also between geographically and/or organizationally dispersed individuals. In this case, one is dealing with a social fabric that may be comparatively loose and the "binding element" may be relaxed to being employed at the same company. Acknowledging this possible difference the approach of this thesis addresses it when investigating what might surrogate the effects of group cohesion such as organizational culture (addressed in section 4.2.5), convincing leadership (addressed in section 4.2.4) and organizational setup (addressed in section 4.2.6) as well as forms of motivation (addressed in section 4.2.11). The thesis also looks at and incorporates theories in the design of the approach that describe loosely-knit social fabrics that act similar to groups as described by Wegner, e.g. the *social capital theory* in section 5.2.1.

The core aspect of Wegner's model, the analysis of distributed knowledge and its effects on individual and group behavior are in line with this thesis' goal: foster the access to knowledge "stored" in the minds of distributed individual. However, the three processes – encoding, storage, and retrieval of information items – are not equally important to this thesis' approach. During transactive encoding two aspects are relevant. First, there are rules in place or are spontaneously established to determine who is responsible to acquire and store the new information. In this thesis' approach, this aspect of extending the overall knowledge base of the organization is out of scope, as the focus lies on the distribution of knowledge within the company that is already available. Nevertheless, the basic theories on how responsibility may be determined, i.e. having shown expertise in a topic, or having taken care of this topic, will inform the design of the mechanism to find suitable interaction partners (see section 5.5). The second aspect during encoding is becoming aware of who knows what. This is one of the major challenges in social fabrics that are rather loose, due to organizational or geographical distance and hence poses great potential for improvement. One of the predominant goals of this thesis' approach is hence to support the acquisition of know-who knowledge. This is also why existing approaches that improve the acquisition of know-who knowledge are analyzed in chapter 3. The storage process is not the focus of this thesis' approach, as it deals with how individuals store information – a field of study that is served by psychology and that only indirectly contributes to this thesis' field of research. Finally, the retrieval of information is of high importance to the approach developed in this thesis. Methods will be proposed that support knowledge transfer without having to live with the bottleneck of *encoding* that limits the scope of potential interaction partners. This thesis' approach will also take care to include a set of interaction partners to allow for transactive enrichment of a retrieval request as suggested by Wegner, to ultimately allow transferring knowledge in mutual direction that is more than "stored" in each single individual (see section 5.5).

This thesis' approach also deals with the challenges of TMS. The bottleneck process *encoding* is specifically altered in the approach to be automated and hence scales to larger social fabrics. Also the employed method for determining levels of expertise, while still not being verifiably correct, provides empirical support for someone's level of expertise that at least gives a good indication. Additionally, the approach addresses the challenge that arises when a TMS dissolves, by incorporating a storage mechanism for interactions. Finally, while not removing the challenge to find a proper design, the approach is designed to allow for the evolvement of different organizational setups. Depending on the configuration of the matching mechanism that may be employed, either redundancy of knowledge or integration of different knowledge is fostered.

2.4 Summary

This chapter established the basis for the remainder of this thesis by first investigating fundamental terminology, such as *knowledge* and *KM* before looking at a sociological theory – the TMS theory – that forms the basis for the approaches looked at in the next chapter and that also is the frame for this thesis.

The analysis of the term knowledge first took a philosophical perspective. It showed that a commonly accepted, "perfect" definition has not been found so far, as with each definition hypothetical examples can be found that contradict the proper definition of the semantics of the word knowledge. In philosophical research, this lead to the question, whether a proper definition is necessary after all and in the scope of the research at hand, this thesis acknowledge the fact that there is no perfect definition. This is why this thesis' pragmatically assume the old definition of Platon, that knowledge is *true, justified belief*, i.e. that it needs to be an articulation about the real world, that really applies and that is backed by some suitable proof. In a further step towards formalizing the term knowledge, different distinctions of knowledge types were inspected. The most often utilized distinction of knowledge is between *explicit* knowledge, i.e. knowledge that can be or is stored outside the human mind, and *implicit* knowledge, that is stored only in the human mind and that cannot be externalized, or if so, only with very high effort. In this thesis' approach the focus will predominantly be on the implicit end of the spectrum, while still acknowledging that some authors argued that the distinction is flawed and both aspects always apply. Further, *hierarchical models* of knowledge have been inspected that lend the semantics of other terms for the definition of what knowledge is. While there are models that differentiate further layers, all acknowledge that knowledge is the understanding of mechanisms that produced information. Finally, to better understand the semantics of the term knowledge different categories were inspected. One typically distinguishes between *know-what*, i.e. knowledge about facts, *know-why*, i.e. knowledge about explanatory models, *know-who*, i.e. knowledge about who knows what and *know-how*, i.e. knowledge about how to do something. Know-who knowledge is what this thesis' approach tries to provide, supporting the possibility to transfer the other kinds of knowledge.

The next part of this chapter dealt with describing the term *knowledge management* by collecting different definitions of the term. As the research discipline is rooted in various parent disciplines there are a number of definitions of which naturally the ones from the information systems discipline are most influential for this thesis. After discussing different attempts within this thesis Krcmar's definition is adopted. It emphasizes the use of *innovative IT* to alter organizational setups for the improvement of knowledge utilization in a company. Then, the evolution of the KM research field was discussed to better understand its heritage and the current problems, which also highlights where new contribution to the field are most beneficial. Three waves of knowledge management can roughly be sketched. In the first, approaches were implemented that attempted to *codify* all knowledge and hence to make it reusable. After some failures, in the second wave the focus was on *connecting individuals* as carriers of knowledge directly. In the third wave, the focus now is on *creating networks of individuals* instead of dyadic relationships. The approach developed in this thesis is part of the third wave of KM, as it focuses on knowledge transfer between individuals while simultaneously spreading the learning effects to a larger group. Another aspect to understand KM is to analyze the tasks that it tries to support. This also allows categorizing where new contributions, like this thesis, will fit in. The most influential model that describes KM tasks is the Probst/Raub/Romhardt model that consists of eight processes. This thesis' approach supports the KM process *Knowledge Identification*, as it helps to clarify who is knowledgeable concerning a specific task. It also supports the KM process *Knowledge Sharing/Distribution* as the transfer of knowledge is at the core of the approach. However, *Knowledge Retention* is also a fundamental aspect of the approach, as it explicitly tries to proliferate the knowledge contained in one-on-one interactions. Finally, as a result of the facilitated interaction *Knowledge Development* is a result of this thesis' approach – sharing of knowledge among individuals leads to increased levels of knowledge within the social fabric, i.e. the organization. Finally, the *SECI* model is introduced and discussed as it is arguably the most influential model to date in KM research. It builds upon the distinction of implicit and explicit knowledge and tries to explain organizational learning processes. According to the model, organizational learning always starts in the tacit knowledge residing in individuals' minds that may be transferred through *socialization*, i.e. model learning. Using metaphors and analogies the tacit knowledge may be *explicated*, explicated knowledge may be *combined* to create new knowledge that then can be *internalized*, i.e. learned. This cycle has, however, received major critique, as explication of implicit knowledge is arguably not possible for all kinds of knowledge. Using the SECI model to classify this thesis' research contribution, the developed approach supports the socialization process, as it is the goal to connect individuals for knowledge transfer and the externalization process, as the results of these interactions are stored.

The third part of this chapter analyzed a theory that tries to describe the processes that underlie the workings of a group with respect to knowledge distribution: the *Transactive Memory Systems* (TMS) theory. Coming from the analysis of intimate relationships it extends the findings on groups to predict group and individual behavior, when knowledge is "stored" in other individuals and accessed via meta-knowledge about the location within these individuals. Building upon this possibility of external storage of knowledge, the theory extends individual memory processes, namely *encoding*, *storage* and *retrieval* of information items onto the group level. In transactive encoding new knowledge items are stored by selected individuals while the others learn about the know-who, i.e. who has the task to store the knowledge item. The rules that determine the responsibility are established already or are setup according to circumstantial properties. The storage process refers to the actual storage within selected individuals, while the retrieval is the process of finding information stored in the set of individuals in the group. This transactive process may also cascade if the requested individual only knows an indirection on the path to the actual knowledge carrier and it may happen that in the communicative acts during retrieval, recoding and amending of information results in new information that otherwise was not possible. While implicitly assuming some cohesion of a group the theory still applies to this thesis' setting, where cohesion may be much lower and the "binding element" may be relaxed to being part of the same organization. Acknowledging this possible difference, this thesis addresses it when investigating success factors for KM approaches in chapter 4, which later serve as one source of requirements for the developed approach. In addition, this thesis acknowledges that the three processes that form the basis of TMS theory, are not equally important to the developed approach. The encoding being automated to a substantial degree allows better scaling but the rules of determining responsibility serve only as inspiration for the design of the algorithmic part in the developed approach. The retrieval will however receive much attention in this thesis' approach, taking care to include a set of interaction partners to allow for transactive enrichment of a retrieval request. In the TMS sub-chapter possible challenges of the TMS were also discussed and they will consequently be addressed during the development of this thesis' approach. The *encoding* process is a bottleneck that limits scalability which is why this challenge is tackled by an automated approach. Further, the challenge that the knowledge within the group is of blurry quality is addressed by evaluating different methods for determination of suitable interaction partners. Additionally, a way to guard against the negative effect that occurs when TMS dissolve is incorporated into this thesis' approach. The loss of knowledge and knowledge retrieval mechanisms is limited by incorporating a storage mechanism for interactions. Finally, acknowledging the results that different organizational goals require different setups of TMS, will inform this thesis' approach to be flexible with respect to organizational goals. Depending on the configuration of the matching mechanism, either redundancy of knowledge or integration of different knowledge may be fostered.

Having laid the fundament by analyzing the relevant terminology and having addressed a theory that aims at describing processes when distributed knowledge needs to be accessed, the next chapter analysis current approaches that try to support the knowledge exchange.

3 State of the Art in Expert Mediation

This chapter analysis the current state of the art in expert mediation and sheds light on their principle ideas, advantages, and disadvantages. This, on the one hand, discloses where there are still unresolved issues and, on the other hand, gives indication for what is working well and is therefore a potential aspect for adopt in a future solution. First, a short overview of possible supportive technology and approaches for KM in general is given, before the focus shifts to those that deal with expert mediation. Three general approaches are identified that have been and are used for bringing together individuals for knowledge exchange, namely *yellow page systems*, *expert recommender systems*, and *knowledge communities*. Each section concludes with a summary of advantages and disadvantages. In a summary, the need for a new approach for connecting individuals for knowledge exchange is motivated.

3.1 Overview of KM Applications and Technologies

Numerous attempts for managing knowledge can be found in research literature, manager magazines, or company reports. Those attempts employ different approaches, applications and technologies, which ultimately leads to a plethora of possibilities that is hard to keep an overview on. This is why there have been endeavors to structure the KM realm. The approach developed in this thesis should consequently be included in existing categorizations of the KM field. Böhmann and Krcmar (2002) provide one such categorization schema. They indicate that IT is a necessary enabler for the distributed setting that is common nowadays and hence focus their structuring on approaches that build upon IT. The technologies are structured according to a two dimensional matrix as depicted in Figure 3-1.

The horizontal axis represents the four phases of knowledge transfer as outlined in the *SECI* model (see section 2.2.4): socialization, explication, combination, and internalization. The vertical axis indicates categories that Böhmann and Krcmar have identified as clusters of technologies, namely archives (or libraries), cartography, team and community support, and knowledge flow. Böhmann and Krcmar shortly describe their categories by indicating their main purpose. *Archive systems* support the storage and retrieval of explicated knowledge. *Cartography systems* support their users by making knowledge better comprehensible for example through visualization or automatic tagging. *Team and community support systems* support the coordination, collaboration, and cooperation of individuals enabling them to overcome spatial or temporal limits. Systems in the category *Knowledge Flow* offer the ability to distribute knowledge similar to the logistics principle: the *right* knowledge to the *right* person at the *right* time. As can be seen in Figure 3-1, most technologies cannot be assigned to only one quadrant of their matrix, because they often support more than one knowledge transformation process or can contribute to different categories of intervention in KM.

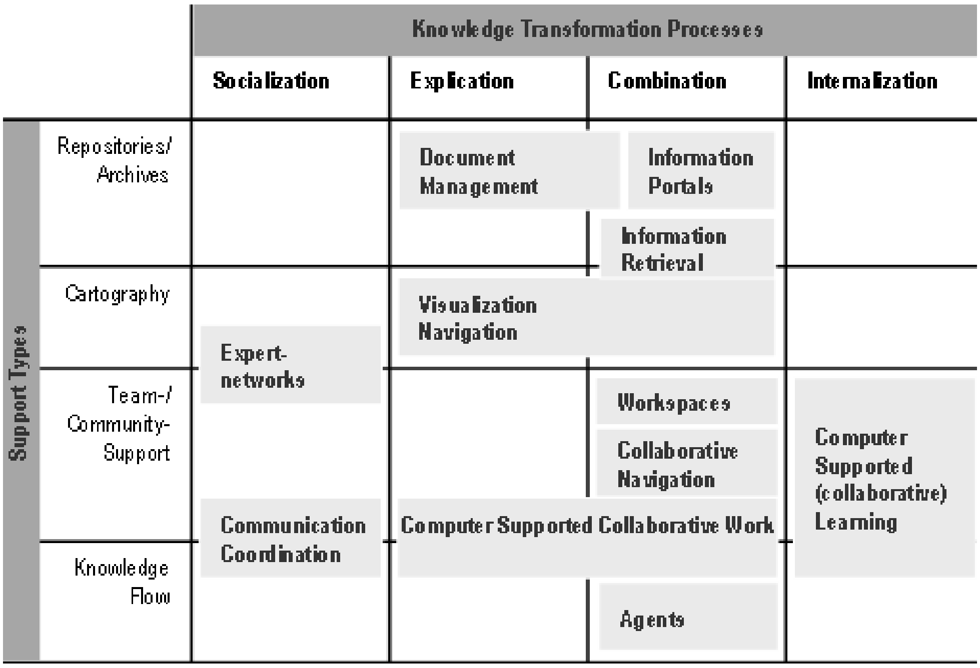


Figure 3-1 ***Technologies for Knowledge Management***
Source: Adapted from (Böhmann, Krcmar 2002)

Another approach for structuring applications in the realm of KM is Binney's KM spectrum (Binney 2001). Reviewing some sixteen sources of reports for KM applications, Binney found that, while sometimes using different terminology, the applications could be summarized in six categories that each cluster similar ideas or areas of challenges that are addressed. He calls those clusters *elements* of the KM spectrum. Figure 3-2 shows the six elements, namely transactional, analytical, asset management, process, developmental, and innovation and creation, along with the applications from the source literature that belong to each of the elements.

	Transactional	Analytical	Asset Management	Process	Developmental	Innovation and Creation
Knowledge management applications	• Case Based Reasoning (CBR) • Help Desk App. • Customer Service App. • Order Entry App. • Service Agent Support App.	• Data Warehousing • Data Mining • Business Intelligence • Mgmt. Information Systems • Decision Support Systems • Customer Relationship Management (CRM) • Competitive Intelligence	• Intellectual Property • Document Management • Knowledge Valuation • Knowledge Repositories • Content Management	• TQM • Benchmarking • Best practices • Quality Management • Business Process (Re)Engineering • Process Improvement • Process Automation • Lessons Learned • Methodology • SEI/CMM, ISO9xxx, Six Sigma	• Skills Development • Staff Competencies • Learning • Teaching • Training	• Communities • Collaboration • Discussion Forums • Networking • Virtual Teams • R&D • Multi-disciplined Teams

Figure 3-2 ***Elements of the KM spectrum and the applications that they host***
Source: (Binney 2001)

Binney describes the six categories he identified as follows. *Transactional KM* applications integrate knowledge directly into operative processes, for example by showing related items while the user enters transaction data. He notes that, while acting upon the presented knowledge is up to the user, the presentation of the knowledge is automatically done by the system whether or not the user wants this. *Analytical KM* applications support the user by uncovering knowledge that is otherwise hidden to the user in vast amounts of information and data. *Asset Management KM* focuses on two things, the management of explicit knowledge, and the management of intellectual property. In this element, any support to make captured knowledge available to users can be subsumed. Binney states that the analogy to a library seems appropriate. *Process-based KM* covers codification and improvement of processes, procedures, methodologies, and work-practices, i.e. knowledge about how things are done. *Developmental KM* focuses on extending the abilities of individuals by offering explicit knowledge in some way or by facilitating the exchange of implicit knowledge between knowledge carriers. The final element *Innovation and Creation KM* focuses on applications that support the exchange of thoughts between individuals, most often of varying background, to generate new ideas, by creating a suitable environment for just this purpose. Binney notes that some of the applications would fit in more than one element of the KM spectrum, and if so, they were placed in the one that was most often associated with the application in the reviewed literature.

Obs	Transactional	Analytical	Asset Management	Process	Developmental	Innovation and Creation
1	Technologist			Organizational Theorist		
2	Explicit			Tacit		
3	Low Optionality			High Optionality		
4	Single Modality			Multiple Modality		
5	Technical Mousetrap			Cultural Change		
6	Structural Capital				Human Capital	
6	Internal Structure				Employee Competence	Internal Structure

Figure 3-3 Trends along the KM spectrum
Source: (Binney 2001)

While compiling the KM Spectrum Binney noted some trends and patterns in the reviewed literature, that better explain the name spectrum that he used for his categorization. He found six dichotomies as depicted in Figure 3-3 that span over the six identified elements. First, he found that each contribution could be assigned to the author's field of previous expertise quite clearly, i.e. management researchers focused their attention more on management aspects, and researchers that are more technology-oriented focused their attention more on technological aspects. In tendency, authors that are technologists cover elements on the left end, authors that are organizational theorists cover the right end of the spectrum. Second, he notes that the KM spectrum's elements focus on tacit knowledge increasingly from left to right. While transactional KM applications almost exclusively handle explicit knowledge, in innovation and creation KM applications the tacit dimension of knowledge is most prominent. Third, he also notes that the degree of free choice to participate increases from left to right. While transactional KM applications mandate the display of knowledge during transaction processing, in innovation and creation KM applications participation cannot be forced and if it is, the quality of the result will likely be low. Fourth, Binney finds that on the left end of the spectrum there is only very limited possibility to use the application in another than the intended way while on the right end of the spectrum different usage patterns are conceivable and the actual usage will depend on how the users deem it appropriate. This is in line with other results from e.g. groupware research (Orlikowski, Hoffman 1997). Fifth, the importance of organizational change and cultural aspects increases from left to right. This seems to be in logical connection to the increase in deliberateness of use and, as Binney states, might be related to the researchers who adopt their known ways of introducing innovations from their respective fields. Sixth, Binney notes that the investments in the applications generate different kinds of values for a company. He utilizes two different categorizations to illustrate this, one that distinguishes between structural capital and human capital and another one that distinguishes between internal structure and employee competence. The elements on the left end increase the intellectual capital more by increasing structural aspects while on the right end the human capital is increased.

In the following, the discussion will focus on a selected set of applications that support the exchange of knowledge between individuals, i.e. on the right end of Binney's spectrum. This has various reasons. First, as indicated in the motivation of this thesis, there is increasingly the perception in research and practice that the management of tacit knowledge, and more specifically the management of individuals and their interactions for knowledge sharing, delivers high value. Most of what companies know is "stored" in the heads of their employees. Second, the applications on the left end of the spectrum have been researched intensively over the last decades. Certainly, there can be optimizations made, but the general approaches in these elements stayed the same over years. As an example, case-based reasoning, an approach to reuse knowledge while catering for flexibility, is still a field of active research. The general approach to inform an individual during the execution of a new "case" by finding a solution of a previous case that has a similar problem statement did not change since the first conceptual idea of case-based reasoning. What did change is the fine-tuning of algorithms and the application of the paradigm to new application scenarios. This is why this thesis is guided by the feeling that more substantial contribution can be gained by addressing the less researched, right end of the spectrum, although some substantial efforts have been made here over the last decade as well. Third, it has become apparent that KM approaches that focus on technology too much rarely work. Rather the focus should be on an integrated approach that takes the individuals' capabilities and motivation and extends those by means of appropriate technology, which also rather points to the right end of the spectrum.

Relating to the KM spectrum this thesis, therefore, focuses on the elements innovation and creation possibly overlapping into the element of development, as learning might take place in discussions that are facilitated by this thesis' approach. Referring to Böhmann and Krcmar's categorization, this roughly relates to a focus on the knowledge transformations socialization with a possible overlap to internalization. On the support type axis of Böhmann and Krcmar's categorization this thesis' output can be located in the area of community support, knowledge flows and categorization. Using Böhmann and Krcmar's nomenclature, this implicates that a focus on expert networks and technology for the support of communication and coordination is suitable. Those two technologies are also applications in the KM spectrum's elements Innovation and Creation, which additionally renders them appropriate candidates for an analysis of the state of the art.

Furthermore, Krcmar (2010, pp. 640) notes, that there are three general processes in KM that need to be managed and that are often addressed by companies that engage in KM initiatives: Knowledge repositories, expert directories and knowledge communities. Knowledge repositories are out of the scope for this thesis as they focus on the storage of knowledge alone. Expert directories facilitate coordination and communication among individuals and may support innovation and creation by increasing the ability to network. Hence, they are eligible technologies to investigate according to the two categorizations described before. In the following two instances of expert directories – Y*ellow Page Systems* and *expert recommender systems* – will be examined. Krcmar further noted knowledge communities to be relevant in organizations. They also facilitate coordination and communication and are often utilized to foster creation of new knowledge through interaction of individuals. Consequently, knowledge communities will be examined in this chapter as well.

3.2 Yellow Page Systems

This sub-chapter first shortly introduces Yellow Page Systems (YPS) on an abstract level before explaining in greater detail the concept of YPS. Having analyzed the concept, the advantages that YPS may deliver are indicated and light is shed onto their disadvantages. The sub-chapter ends with a summary of findings that inform the design of approaches that specifically try to address the downsides of YPS.

Most of the knowledge in a company is not available in digital or otherwise explicated format but is only "stored" in the heads of its employees (Krcmar 2010, p. 663). This tacit knowledge cannot simply be searched for and directly retrieved as it is possible for explicit knowledge stored in documents. To “retrieve” this knowledge it is necessary to communicate with the knowledge bearer to transfer the knowledge in a social, communicative process. A prerequisite for this is having the necessary *know-who*, i.e. to know who is in possession of tacit knowledge that is currently needed. This is where YPS come into play. These systems that are often also referred to as expert directories or expert maps, aim at making non-codified knowledge accessible by allowing a knowledge seeker to identify an individual that holds the desired knowledge.

3.2.1 Concept

The name Yellow Pages is adopted from the similar concept of a specific type of telephone directory. There, the idea was to have access to a list of companies and freelancers that are not sorted by name but by branch or profession. Likewise, a YPS in KM should allow a knowledge seeker to query for a piece or category of knowledge and to find those individuals who possess the desired piece of knowledge. If a YPS is in place, the knowledge seeker can enter a search term and the system will evaluate which of the employees have competencies, skills, etc. that match the entered request for knowledge. The ones that have a match are shown to the requester. If the individual profiles not only contain knowledge items but also distinguish levels of proficiency for knowledge items, it is also possible to show a ranked list to the user searching for specific knowledge. To be effective, each employee page includes contact details, such as phone number or e-mail address allowing the knowledge seeker to step into contact with the knowledge bearer. Additional information, such as organizational position or socio-metric information like age or gender may also be included in the profile pages, to increase a feeling of connectedness between the knowledge seeker and knowledge bearer and to allow the knowledge seeker a selection of a contact person with which he feels most comfortable.

To setup a YPS there are some steps to be performed in advance. They capture skills and competencies that comprise knowledge items or categories. Therefore, as a first step, a company implementing a YPS must determine which competencies are relevant for supporting its operations. Those competencies will be most relevant that are crucial for the core processes of the company. Also those that will be searched for most often are most relevant. For example, a relevant competency could be “proficiency in programming Java”, but also a less tangible knowledge item such as “experience in leading software development teams” could be relevant for a company. This shows one of the challenges in defining competencies. It must be apparent which knowledge items are relevant and also on which level of abstraction. For example, the two competencies mentioned before could have been subsumed by only one very abstract competency “Knowledge of software development” or, on the contrary, the competencies could have been much more fine-grained. Therefore, instead of “proficiency in programming Java" a company might instead use "proficiency in user interface programming using Java" and "proficiency in client-/server programming using Java". Obviously, the suitable balance between only some coarse competencies and many very specific ones depends on the specifics of the company but may not be easy to identify. Additionally, it might be necessary to distinguish between hard skills or *qualification knowledge* such as the proficiency in a programming language and soft skills or *experience knowledge* such as the ability to work in teams (Gronau, Uslar 2004; Deiters et al. 1999).

In a second step, the company must think of the levels of expertise that should be used. Relating to a competency "Customer Service using SAP" a company might want to distinguish four levels of expertise such as "basic", "working", "leadership" and "expert" (Davenport 1997, p. 205), while another company rather uses a numeric five-point scale. Again, it depends on the unique needs of the company. The adoption of the levels among the employees will be relevant. If it is not apparent to the employees, what a competence level of four out of five means, natural language expressions might be better. In a third step, the company must determine the introduction procedure. It is necessary to determine how the profiles should be populated. If a company uses an electronic directory service, the contact data such as e-mail address and phone number can be easily imported from this source (Becker et al. 2009, p. 646). Further, if a company has a skill management system used for Human Capital Management (HCM purposes, YPS are often integrated with them and can import formal qualifications easily from them (Gronau, Uslar 2004). However, while for formal qualifications the automatic population might be possible, the experience knowledge cannot be extracted from possibly existing systems. Here the question arises, whether employees should judge and rate those competencies themselves, if it should be done by the supervising manager, after performance evaluation meetings jointly by the supervising manager and the employee or by peer judgment from coworkers or maybe even a larger group (Gronau, Uslar 2004; Davenport 1997, p. 206). Especially if employees judge their competencies on their own, they must possess the ability and willingness to be self-reflective to capture their skills and competencies adequately. Likewise, when a supervising manager evaluates the skills and competencies of an individual he needs to be able to reflect the past results of an individual and deduce corresponding levels of expertise.

In a final step, the company has to decide how the YPS can stay up-to-date. Especially in knowledge-intensive, rather creative work settings, knowledge in a company has only a short lifespan (Gronau, Uslar 2004). It changes frequently as it is outdated with respect to the current body of knowledge, is expanded through work experience or formal qualification, or individuals might start to learn new competencies which also needs to be reflected in the YPS. However, the data quality must be high for the acceptance of the system, which is why there need to be organizational and/or technological measures to keep the data up-to-date. This naturally inflicts costs in the form of employees' time (Deiters et al. 1999).

Being a first approach to manage the tacit dimension of knowledge within a company, the use of YPS became popular in the end of the 1990ies. Examples of companies and their respective YPS' names include Microsoft's SPUD (Davenport 1997), HP's Coonex (Davenport 1997), a competency database implemented in a PeopleSoft solution at Ernst&Young (Davenport 1997), and Teltech's list of external experts along with their expertise that consists of keywords from the company's thesaurus (Davenport 1997).

3.2.2 Advantages

Well-maintained YPS allow finding employees using only search terms that correspond to desired qualifications or skills. If the data is of high quality and the competence and skill catalogue contains terms that are familiar to the user, the timesaving for finding potentially suitable colleagues can be great. One additional advantage of having a YPS in place is its potential to support other tasks besides offering support for finding an individual with knowledge that someone currently needs. YPS can also support HCM (Gronau, Uslar 2004). If used for this additional purpose, they are typically integrated into HCM systems and exchange information with professional development activities. Containing skill and expertise information, YPS may also support staffing of projects. Once it is clear which competencies will be necessary for a certain project, the person responsible for staffing can swift through the YPS to find suitable candidates. YPS can likewise support professional development, also called Skill Management. Having an overview, which and how many employees have necessary knowledge it is possible to take the current situation as basis for planning qualification measures to have adequately skilled employees when needed in the future.

3.2.3 Disadvantage

While offering some distinct advantages, YPS also face some severe challenges. This is why the traditional type of YPS as introduced and discussed above often fail after not too long. The challenges are detailed in the following.

First, the initial development of a YPS is very resource-intensive. Often it is not obvious in a first iteration, which competencies should be modeled and elicited, which necessitates further iterations. As all employees have to be queried, this can become very resource-intensive. Limiting it to only a smaller fraction of employees might be less costly, but then the company runs risk to not capture all needs and also the system might fail to be accepted in the first place as it does not contain sufficient amounts of data to be helpful.

Second, the time and willingness of experts to disclose their levels of expertise is crucial. However, there is no obvious benefit of explicating one's knowledge in yellow page systems

(Schütt 2003, p. 458). If everyone in the company receives similar benefit from the system, this might be a motivation for employees to keep their profiles up-to-date – give and take. However, as it is often the case, if there are some that have demanded skills and competencies and many that query those, then this motivation is no longer there (Deiters et al. 1999). In this case, explicating more of one's competencies will result in more work which is certainly rather a reason to not indicate one's expertise. Other organizational measures would be necessary in this case. Additionally, the problem of keeping the contained information up-to-date also applies for the catalogue of competencies and skills itself. Models of expertise are typically coarse while retrieval of experts typically focuses on very fine-grained, specific needs (Seid, Kobsa 2003, p. 5). Therefore, the models need to be adapted frequently themselves.

Third, changes in expertise level are frequent due to constant changes in organizational life. Individuals might extend their degree of profession in one area, might become knowledgeable in another, or might degrade in proficiency relative to other coworkers. To reflect these changes, each employee would have to keep their profiles up-to-date, which is often not the case, as it is time-intensive and labor-intensive for the individual himself. This is especially critical since lack of time is seen as the major inhibitor to KM in general (Frauenhofer IAO und Deutsche Bank 1999; KPMG 2000). Therefore, the databases are quickly outdated and keeping them up-to-date is again very resource-intensive and often too burdensome for the experts.

Fourth, YPS are typically systems on their own. From a technical point of view, they might be integrated into skill management systems or directory services for more convenient import of data and possibly also to reflect data back into e.g. the skill management system. However, they are not integrated into the operational systems that individuals use for their actual work. On the one part, this means there is a change in media or system used, which might be an additional burden for the user. On the other part, this means that the knowledge seeking individual must know what he needs to know, i.e. he must be able to articulate his need for an expert by realizing which competence or skill he currently lacks. Without integration into an operational system, there is no way to support this mental process as the context of the individual cannot be elicited.

Fifth, the intention of YPS is the facilitation of the socialization process (see 2.2.4). However, the focus is on one-to-one interaction, i.e. one individual uses the YPS to find possibly many suitable candidates for querying but usually only queries one. This leads to a relatively unmanaged network of knowledge flows between individuals (Oldigs-Kerber 2007, p. 67), where the fruits of the interaction are neither stored electronically in some form, nor in the heads of others but the two interaction partners.

The problems of yellow page systems are summarized in Table 3-1.

Id	Description
Problem 1	Initial setup is resource-intensive as modeling is not straight forward
Problem 2	Too high effort for the individual to keep data up-to-date; low motivation of individual
Problem 3	Frequent changes incur either high cost or information is not up-to-date
Problem 4	Non-integrated; lack of work context
Problem 5	Creates one-on-one interactions, that are not well manageable

Table 3-1 ***Summary of problems that YPS face***
Source: Own Illustration

3.2.4 Summary

YPS try to support transfer of knowledge through mediating the socialization processes. This is achieved by enabling an employee to search for necessary competencies similar to searching for documents. As a prerequisite, each individual in a company needs to maintain a profile that states his competencies and skills as well as contact information. The company has to setup adequate competence and skill catalogues for this purpose. The initial data entry as well as the continued maintenance of an individual’s profile is done manually, either by the individual alone, or by the HCM department or the supervising manager. However, the individual can hardly be relieved entirely, especially for capturing soft skills that rely on personal experience and not so much on more obvious indications like formal training. Independent of who maintains the profiles, it is either very resource-intensive as changes may be frequent or the information is quickly out-of-date. As experience from the first documented research on real-life applications showed, the approach turns out to be inadequate in most cases.

When creating more suitable KM applications some learning can be deduced. First, the effort for the individual needs to be reduced considerably. If too much time for maintenance of a profile is one of the major reasons for failure of the approach, as has been stated, then exactly this effort and the time necessary for setup and maintenance of a profile needs to be reduced. Second, the ability and motivation for self-reflection cannot be taken for granted. Individuals may know more than they think or on the contrary, they might, knowingly or accidently, have entered exaggerated levels of experience. Therefore, the data in the YPS cannot be trusted in the sense that there is little tangible evidence for the claims. A more advanced approach should try to provide evidence for competence and skill statements to move them away from being subjective assumptions. Third, suitably modeling competencies and skills is not easy and the right level of abstraction is difficult to find. Additionally, individuals search for expertise in specific situations and therefore have a very contextualized demand. This results in the need for very fine-grained searches – often more fine grained than the competencies model may support. In line with high effort for keeping individual records up-to-date, the underlying schema also needs to be updated and very fine-grained. Therefore, there must be either technical or organizational means to keep the model of competencies up-to-date. Fourth, the individual should be supported when searching by allowing the context to influence the search and this way to reduce the mental effort. Fifth, there should be means to generate a learning for the company as a whole in some way in addition to the learning that takes place between the two interaction partners.

The drawn conclusions are summarized in Table 3-2.

Id	Description
Conclusion 1	Effort for the individual must be reduced as much as possible
Conclusion 2	Proof for competency or skill should amend or replace self-reflection
Conclusion 3	Means for supporting the maintenance of competency models is necessary
Conclusion 4	The current context of an individual should be taken into account
Conclusion 5	Means to increase the benefit from the interaction to more than just the two interaction partners are desirable

Table 3-2 ***Summary of conclusions from analyzing YPS***
Source: Own Illustration

3.3 Expert Recommender Systems

This sub-chapter first shortly introduces Expert Recommender Systems (ERS) on an abstract level before shedding some light on recommender systems in general as they form the basis for ERS. Then, in greater detail, the concept of ERS is explained. Having analyzed the concept, then the advantages that expert recommender systems may deliver are indicated and light is shed on their disadvantages. The sub-chapter ends with a summary of findings that inform the design of approaches that specifically try to address the downsides of ERS.

The general idea of an ERS is to remedy the problems of traditional YPS and to take the concept one step further by automating tasks such as the profile generation, the mediation of communication and others. In the course of expert seeking, the system should address the user’s direct and indirect questions. Generally, when someone seeks an expert, he wants to know *if* there is an expert that can answer the user's questions, what *level of expertise* the user has and *how it compares* to others, if *there are others* that also fulfill the criteria and *how* the person can be *reached* (Seid, Kobsa 2003, p. 5).

Having created a substantial field of active research, it is not surprising that the concept of ERS has been labeled with different names that all refer to the same idea. In accounts of research and practical applications, the following terms are used synonymously:

- Expert finding system
- Expertise finding systems

- Expert search
- Expert recommendation
- Expertise location system

In the following, the term Expert Recommendation System (ERS) is used, acknowledging that some cited authors have used one of the synonymous terms instead.

ERS are an instantiation of the more general class of recommender systems. To better understand these systems, therefore an investigation of the more general area of recommender systems is reasonable. After a brief introduction, the focus will be on the principle ways of eliciting recommendations and how those may be interpreted in the light of ERS.

3.3.1 Recommender Systems

Recommender systems are the focus of an established field of research that started to form in the middle of the 1990ies (Adomavicius, Tuzhilin 2005), with (Resnick, Varian 1997) representing an often cited, fundamental contribution to the field. The main goal is to reduce the problem that comes with *information overload*. The information sources available to users offer much more items than the user can inspect, also reflected in a common saying: "We are drowning in information but starved for knowledge" (Naisbitt 1982). Recommender systems try to reduce this problem by proposing those items out of the large set of possible ones that offer most utility to the end user. The utility of an item can be elicited by different means, i.e. there could be commercially-motivated rankings of items that are presented to the user or explicit user ratings are used as basis for utility functions. Hence, the recommendations are based on ratings of users for certain items that were previously used or looked at to deduce ratings for items that are yet unknown to the user. The items that are likely to have a high personal rating with respect to the rating history of this user are then recommended to him. The research field evolved over time focusing on the optimization of the algorithms that compute the recommendations. The kinds of algorithms and therefore the concepts however remained constant over the years. In the following, first a classification of recommender systems is presented and then different kinds of recommender systems are explained.

3.3.1.1 Classification Schema for Recommender Systems

The field of research in recommender systems has expanded over the years, as have the prototypical implementations and the nuanced improvements that could be obtained. To get a better understanding of the broad field, there have been attempts to structure the research field. In the following, a suitable structure is adopted and its classes are analyzed.

Recommender systems are often classified according to the principle way that they employ to determine suitable recommendations. The following three classes are often distinguished (Balabanović, Shoham 1997):

- Content-based recommendations
- Collaborative recommendations
- Hybrid approaches

Content-based recommenders use models about the items to be recommended as their basis, *collaborative recommenders* use models about the users to recommend potentially relevant items and *hybrid approaches* utilize combinations of these two classes. Burke (2000) distinguishes another class: *knowledge-based recommender* applications, that will be discussed as well. Terveen and Hill (2002), on the other hand, separate four somewhat different classes of recommender systems:

- Content-based recommender systems
- Collaborative filtering recommender systems
- Recommendation support systems
- Social data mining systems

The first two are in line with the previous approaches to classify recommender systems. *Recommendation support* is not a recommender in the narrower sense but offers only a platform. It allows users to either rate items or annotate those items with free text. Other users can use these rating to filter items, i.e. only take those that are rated above a threshold value (Terveen, Hill 2002). This approach works fine if filtering on good or bad is useful, i.e. if the domain itself is already known and the problem lies in filtering within that domain or ranking within it. However, in expert mediation the domain of relevant experts is not known, therefore filtering out is not the prominent challenge that needs to be addressed. This kind of recommender system is therefore not further inspected. The final class of Terveen and Hill are *social data mining systems*. They classify every recommender system that extracts implicit preference information from typical usage situations as social recommender systems. The relevant computational records can be documents, Usenet messages, or web sites and links. (Terveen, Hill 2002). This is quite relevant for every application that tries to be less obtrusive in determining the user's preferences. However, this class does not appear to be mutually exclusive as many recommender systems are using either collaborative filtering or content based approaches but extracting at least some features automatically. Therefore the further discussion relies on the classification of Balabanović and Shoham (1997) extended with the knowledge-based recommender class.

The approaches outlined in the following differ with respect to their demands upon and usage of three artifacts: the item universe, the user universe, and the utility function. The differences will be illustrated in figures in the following. The *item universe* contains all possible items including both the ones that a user has interacted with before and the ones that he has not interacted with before and that are candidates for recommendation. Items can be described with different levels of detail. The *user universe* contains all users that interact with the recommender system, i.e. the individual and his peers. The *utility function* projects how "valuable" a recommendation will probably be for a user and how valuable items are that the user has interacted with before. Various kinds of information can be used for this purpose.

3.3.1.2 Content-Based Recommender Systems

Approaches in this class of recommender systems can be seen as a natural extension of classical Information Retrieval to the application field of recommendation. Systems recommend those items that are similar to the ones that the user preferred in the past. For this to be possible, the items that are to be recommended, i.e. the content, needs to be described by features that allow to determine similarity between items sufficiently well. The general idea is depicted in Figure 3-4. The system determines those items that the user knows and for which he, or an external entity, has assigned high utility values. Using the features of the item universe as ingredients a *similarity measure* determines those items that a user does not know yet and that are most similar to the ones with high utility determined in the step before. The most similar item or a (ordered) list of most similar items is then presented to the user as recommendation.

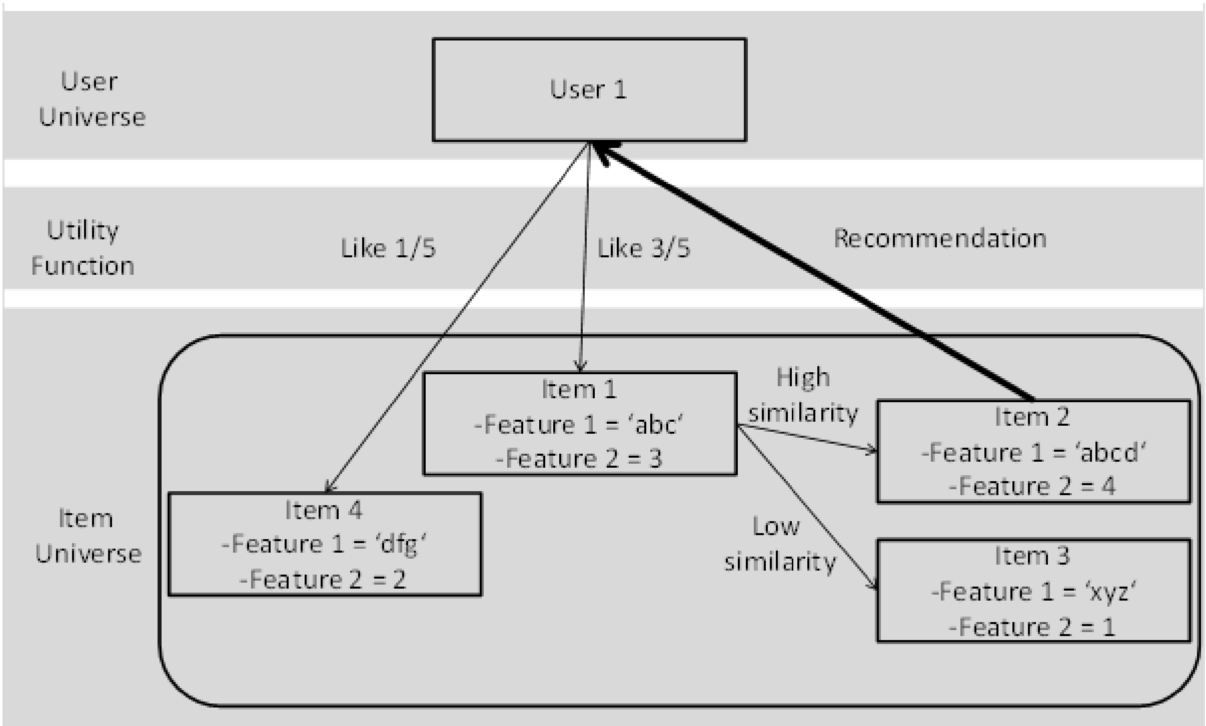


Figure 3-4 ***The concept of content-based recommendation systems***
Source: Own Illustration

The features of the items either have to be maintained manually or are assigned automatically using some form of inference mechanism. Relating to ERS, the items would be individuals and the features would be the expertise of the individual. In classical YPS, the profiles for each individual are maintained manually, in modern ERS the expertise is extracted automatically as discussed later.

The challenge of automated approaches is to ensure that the automatically extracted features have a high correlation with the actual utility. Another challenge lies in the fact that the user is only presented with a subset from the item universe consisting of the same kind of items that he knows already, as those form the basis of the recommendation. Therefore, already known kinds of items will be recommended, while new items are not recommended or recommended much less. The concept of content-based filtering is also not directly transferable to expert recommendation. Typical content-based recommendation systems require the user to rate items, which translates to other individuals in ERS. However, when using ERS the knowledge seeker does not know how to rate other individuals with respect to his current need, which is why he uses the ERS in the first place. Therefore, support in terms of the utility function need to be supplied.

3.3.1.3 Collaborative Recommender Systems

Recommender systems within this class, make the assumption that *similar individuals* will like similar items in correspondence to the saying “birds of a feather flock together”. The recommendation is produced by first finding individuals that are most similar to the user who queries the recommender system. Then, the items that have high utility for those most similar individuals are presented as recommendation to the initial user. This concept is also depicted in Figure 3-5.

As the similarity between two individuals typically relies on the overlap of ratings in the item universe, collaborative recommender systems face a similar challenge as content-based systems do: Users need to rate items so that the algorithms can work. In ERS this again translates to rating other individuals with respect to the expertise sought – the very problem that ERS try to tackle. Hence, direct transfer of the idea is not possible, but the similarity between users could be determined in a different way to still rely on the principle idea. Additionally, collaborative recommender systems face a challenge with new items in the item universe. If no one has rated the item yet, it is also not rated by similar individuals and will not show up as a recommendation. In ERS this would mean, that a new employee with special knowledge would only be recommended as an expert once someone has indicated that this individual is an expert in this field. Finally, collaborative recommendation has to deal with sparsity. Typically, users rate only a very small fraction of all items in the item universe. Therefore, the possibility for an overlap of ratings is small. Aggregation mechanisms that increase the overlap of users’ ratings or other mechanisms to indicate which users should be considered peers are therefore often indispensable for this approach. In ERS this relates to the problem, that the approach only works if there is a sufficient number of peers that have relied on the same experts before, given that the similarity is determined as in the original field of application, i.e. by overlap of rating.

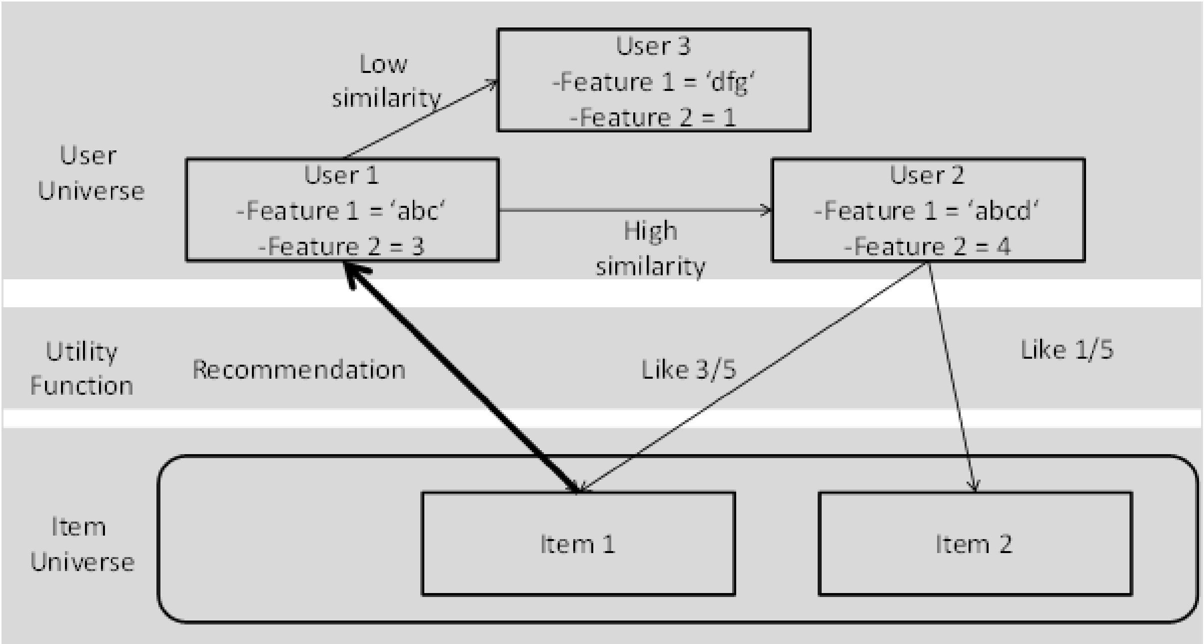


Figure 3-5 ***The concept of collaborative recommendation systems***
Source: Own Illustration

3.3.1.4 Hybrid Recommender Systems

Combining the content-based and the collaborative recommendation approaches, it is the goal to have the advantages of both approaches while trying to limit their disadvantages. Adomavicius (2005) lists four ways to combine the two approaches:

- Using both approaches as they were but combining their predictions
- Integrating some content-based characteristics into the collaborative approach
- Integrating some collaborative characteristics into the content-based approach
- Creating a unifying model that incorporates characteristics of both approaches

Those combined approaches increase accuracy by intelligently exploiting the different approaches' advantages. However, they do not conceptually present new approaches. For the course of this thesis, those combinations are not inspected further, as the focus does not lie on efficiency or effectivity of general recommendation systems.

3.3.1.5 Knowledge-Based Recommender Systems

Recommender systems within this class try to resemble an advisor. The items proposed by the system are not (only) selected by previous preferences of the user or his similar peers, but by constraints and logical dependencies within the item universe that limit the choice of useful items. Hence, this class of systems relies on explicit modeling of a knowledge base. The knowledge base consists of the following parts (Felfernig et al. 2006, pp. 15–16) :

- Information about features of the user. If the user is a customer this could be his upper limits for purchases.
- Information about the items. If the item is a product to be sold, this could be the price or product category.
- Constraints on the combination of user features. If the user is a customer he might for example not have the feature “upper limit for purchase 1000 €” and “affinity for high-price products” at the same time.
- Criteria for filtering based on item and user features. If the user is a customer with the feature “affinity for low price products” and a product belongs to the premium category, the customer should not be provided the product as recommendation.

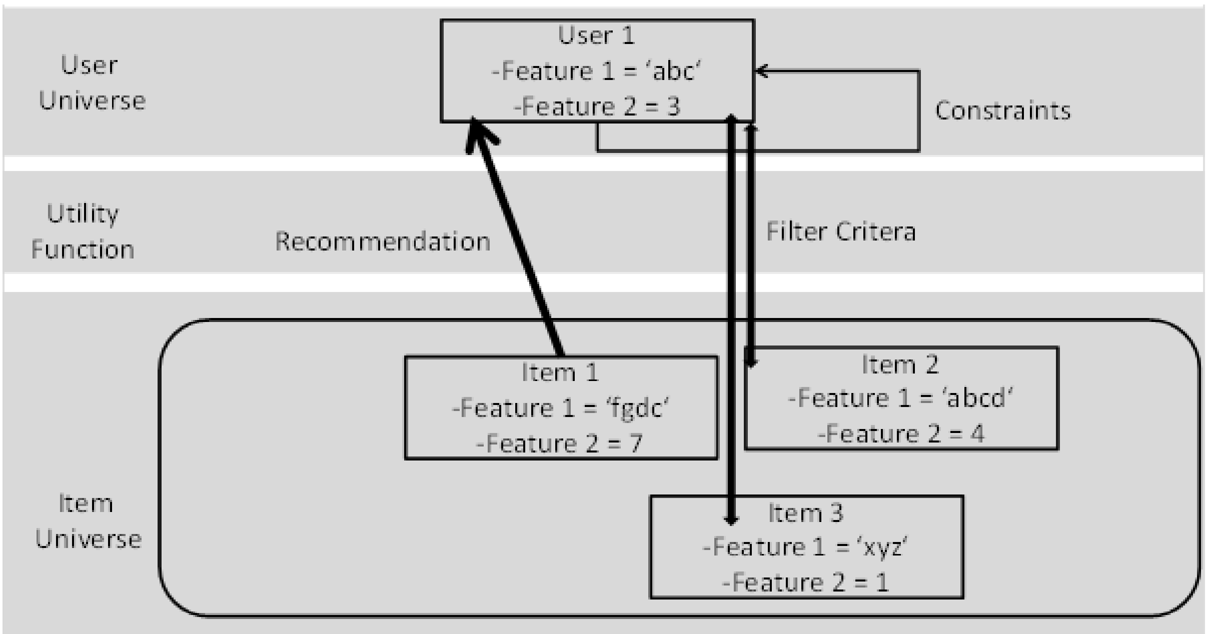


Figure 3-6 ***The concept of knowledge-based recommendation systems***
Source: Own Illustration

The relationship between the components of the knowledge base are shown in Figure 3-6. *Constraints* limit the number of possible configurations of a user model instance. *Filter criteria*, on the other hand, remove items from the list of possible recommendations. The remaining items are ranked using further rule inference mechanisms. The inner workings of those inference mechanisms can depend on rules within the item universe, i.e. relationships between items and deduction rules. It is for example possible, that the interaction of the user is taken into account, i.e. if the user did not like an item during a session then another class of items is presented as recommendations based on a rule that has the user's dislike of the initial item as input information.

Knowledge-based recommender systems are an adequate solution if the selection process of the recommendations needs to be made transparent to the user. Felfernig et al. indicate that collaborative filtering approaches as well as content-based recommender systems are inferior to knowledge-based recommenders when explanations of the recommendations are necessary. The existence of a knowledge base allows far more detailed descriptions (Felfernig et al. 2006, p. 13). This is why recommender systems of this kind have been used for guiding users in their web browsing behavior (Burke et al. 1997) and in recommending complex products such as financial services or high-tech products (Felfernig et al. 2006). Some knowledge-based recommender systems also start conversations with their users for advising them properly (Thompson et al. 2004; Jiang et al. 2005) which is sometimes also referred to as (online) consumer decision support. By offering automatic guidance, knowledge-based recommender systems have some resemblance and share some commonalities with *Expert Systems*, where it is the goal to create applications that can assist users similar to how a human expert would (Darlington 2000).

When transferring the concept to ERS, the knowledge base could be interpreted as follows. The information about features of a user would be information about the preferences and abilities of a knowledge seeker. The information about items would be preferences and possibly further socio-metric information about experts, constraints would similarly ensure the logical correspondence of features within the user model and filter criteria would resemble rules for who should or should not be recommended to users in a certain context or with certain dispositions. This transfer of general concepts to the specific case of expert recommendation appears to be sound. As a matter of fact section 3.3.2.7 will elaborate on a type of expert recommendation that utilizes explicit models of users (user universe) and knowledge (part of the item universe).

3.3.1.6 Applications of Recommender Systems Related to Knowledge Exchange

As discussed, ERS are the subclass of general recommender systems that is relevant for this thesis' context of supporting knowledge exchange. Some other subclasses of recommendation systems pursue different yet related goals. As those related recommendation systems have the potential to inform a suitable solution to the challenge of expert recommendation, they are sketched in the following.

3.3.1.6.1 Recommending Workflow Execution

One relevant challenge in knowledge-intensive work is knowing which steps to perform to reach one's goal in an efficient and effective manner. If a process is well understood, repeats fairly often and exhibits only limited variability, workflow systems can incorporate a model of this process and can support the user by guiding him through each step. If the process is either not well understood yet, is not repeated often enough to bear the cost of implementing a workflow system or if its execution is subject to high variability, then workflow system support is not possible at all or only to a limited extend. Vanderfeesten et al. (2008) report on a workflow execution system, that is more flexible than standard systems and allows to support also processes that cannot or are not modeled entirely. While workflow systems obviously need an execution model, i.e. a process model, Vanderfeesten et al. describe a product-based workflow support system that operates on a product model. A product model illustrates what needs to be done but not in what order and not in what way. By relying on only a product model, there will be, nevertheless, decisions to be made, i.e. which next step in the workflow to execute. The suggested system uses the product model and additional customizable algorithms that optimize on features such as overall runtime, overall cost or number of work handovers. Knowing the current state of the process with respect to the product model, the system recommends a next step in line with the optimization criterion specified in the algorithm.

Recommending next steps in a complex process is certainly a valuable support for knowledge work. This thesis, however, focuses on knowledge transfer from individual to individual which is why the workflow execution recommendation is not directly influential for this thesis. Yet, the idea to utilize the *execution steps* that have been performed already in a knowledge-intensive process appears interesting. This thesis approach will adopt a similar approach by taking into account the previous interactions of the users for finding suitable interaction partners.

3.3.1.6.2 Recommending Web2.0 Contributions to Write

All systems that have been looked at so far, support knowledge transfer in a reactive manner. When an individual needs some specific knowledge the systems support finding a knowledge bearer. This makes sense, as the need to rely on external knowledge from someone else arises in situations where there is actual need, for example for fulfilling a certain task. Another way of looking at knowledge transfer operates in a proactive manner. The focus lies on spreading knowledge to individuals who are not directly in need of it but who want to learn independent of their current context. One such way is encouraging individuals to write blogs, that other individuals subscribe for. Geyer and Dugan (2010) found that this way of knowledge dissemination could further be supported. They used a survey to find out more about the blogging behavior within a company. The results indicated that the usage of blogs, which are often implemented as first tools in companies that want to take advantage of the Web2.0 wave, could be increased, when the blog writers are aware of the need for a certain topic. Geyer and

Dugan interpret blogging as a social interaction and reason that users would be more likely to blog if they knew about the interests of their potential audience. Their survey results further indicated that especially less-seasoned bloggers welcome the recommendation of topics to write about. In addition, they found that contributions are more likely when the request for contributing on a certain topic is directed to them in person and not just posted anonymously.

This thesis' approach of supporting the knowledge exchange will be rooted in the current work context of the user. Learning through push-based knowledge dissemination from other users is therefore outside of its scope, while certainly also an endeavor worth the while. However, Geyer and Dugan's recommendation system and the survey they conducted, brought up some ideas that can be adapted for this thesis' approach. Their finding that especially less-seasoned bloggers need support can be transferred to knowledge exchange where less-experienced users might not feel comfortable to contribute. Also, their finding that requests directly posted to a person instead of anonymously increase the likelihood of generating a contribution is possibly transferable to this thesis' setting where posing a request for knowledge directly to someone as an expert in this area might be responded to more likely then when posting it to an anonymous group.

3.3.1.6.3 Commercial Application of Recommender Systems

Recommender systems have formed an active research field, but likewise the commercial world is also attracted to their potential. There are some prominent examples of the use of recommender systems. Amazon.com uses recommender systems to propose items that might be interesting for a user based on what other similar users liked (Linden et al. 2003). MovieLens offers a similar service for movies (Miller et al.) while there are also implementations for the recommendation of news articles (Billsus et al. 2002).

While the purpose of recommending new products is certainly outside the scope of this thesis, appreciating the fact that there is commercial usage of recommendation system gives a vague indication for two things. First, it indicates that the concept of recommending items indeed offers the potential to deliver benefit by e.g. increasing sales in the above examples. While recommending items that are just "things" is different to recommending experts, which includes social processes and constructs, the same assumption is underlying: Proper recommendations increase the utility of the domain that they are applied on. Second, the use of recommender systems in large scales as in the examples above indicates that the users appreciate the support or at the very least do not dislike it. Again, transferring this result from the world of recommending things to recommending experts is not directly possible. Nevertheless, it still indicates that the very idea of receiving recommendations is likely to be accepted among users.

3.3.2 Concept of ERS

This section inspects the general concept of expert ERS. First, the meaning and constituting steps in expert recommendation are explicated before possible and often utilized sources for expertise determination are described. Then different approaches for expert recommendation are detailed subsequently.

The need to find an expert can arise in two different settings. Either someone needs *special information* from an expert and uses this other expert as external memory or someone needs an expert to *perform a specific organizational or social function* (Seid, Kobsa 2003, p. 3). While the distinction may be fluent, there are some differences depending on which of the two the focus lies. In the first case, that is also more in line with the goal of this thesis, the focus lies in supporting the socialization process (see section 2.2.4) and is hence directed at supporting the transfer of tacit knowledge. In the second case, the focus lies more on a HCM task like staffing teams, which is less the focus of this research. In both cases however, ERS support the identification of experts using traces of explicated knowledge such as documents or e-mails to infer expertise. In a second step, ERS might offer possibilities to initiate the socialization process. As illustrated in Figure 3-7, the services of a comprehensive ERS are conceptually divided into expertise or expert identification and expertise or expert selection (Maybury 2006; Serdyukov et al. 2008).

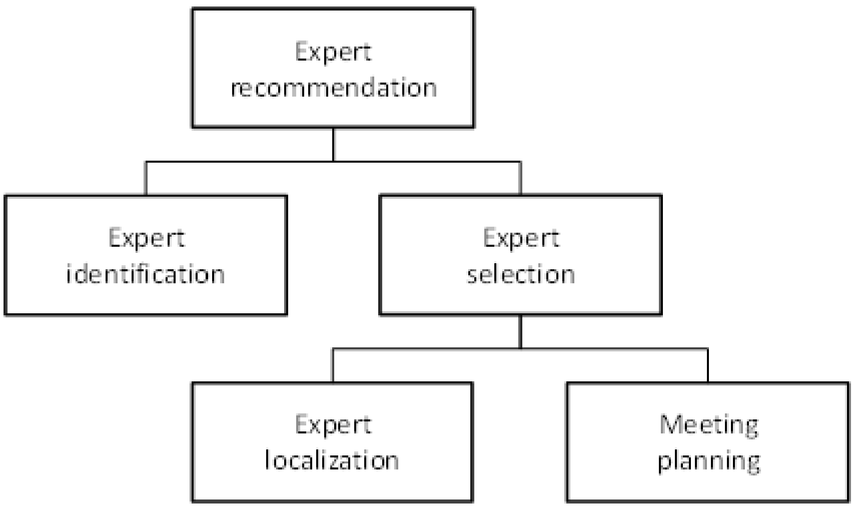


Figure 3-7 Constituting parts of an expert recommender
Source: Own Illustration on the basis of (Maybury 2006; Serdyukov et al. 2008)

The goal of *expert identification* is to take a knowledge seeker's request, which is typically formulated as a query for a certain type of expertise and produce a (possibly ranked) list of those employees that are expected to hold this expertise.

The goal of *expert selection* is to find those experts for which a synchronous meeting is possible and desired by both parties (*expert localization*) and subsequently to schedule such a meeting (*meeting planning*). The physical location of the employees, available time, personal preferences, and possibly company regulations can constrain the possibilities and need to be taken into account during expertise selection.

To better understand the processes that trigger the need to find an expert for special information, Seid and Kobsa (2003) analyzed which circumstances lead someone to consult an expert. First, someone might need access to information that is not documented, because it simply has not been documented yet, it cannot be documented but only be transferred through a social process or it is deliberately not stored in a retrievable format for economic, social or political reasons. Second, someone might not be able to exactly specify what he needs to know, rather the dialogue with an expert is the process that facilitates information retrieval either from this expert, another expert or a knowledge repository. This is in line with TMS theory's transactive retrieval in social groups (see 2.3). Third, someone might want to utilize an expert to be more efficient. Someone with advanced expertise can handle tasks faster than novice users can. Therefore, relying on the expert improves the initial users efficiency. Fourth, often users do not want a context-free, general piece of knowledge but rather need a contextualized, situated interpretation of more general knowledge that the expert might poses. Therefore, the expert acts as the specialization mechanism for more general knowledge that the user might not comprehend. Fifth, someone might simply prefer relying on social interaction instead of using anonymous media like documents (Seid, Kobsa 2003, p. 3).

Before looking at approaches that support situations, in which expert finding is necessary, a prerequisite to this is discussed. In order to determine suitable expertise automatically, sources that can be used to infer expertise need to be found and tapped into. This is why it is necessary to study the different classes of sources that are typically used to infer expertise of individuals.

3.3.2.1 Sources for Automated Expertise Determination

The envisioned advantage of ERS in comparison to YPS is the automatic generation of information about the expertise of its users. The automation is done by relying on tangible evidence that is used to infer expertise. This tangible evidence can be found in different kinds of sources. Heeren (2001) indicated that the following four classes of sources can be used to extract evidence for expertise:

- Communication-related sources (e.g. e-mails or newsgroup postings)
- Document-related sources (e.g. files, documents in a CMS, or websites)
- Appropriated Knowledge Assets (e.g. citations)
- Interaction recordings (e.g. resource usage, program usage)

Communication-related and *document-related* source are used most often in prototypical research systems and commercial ERS, some approaches, however, also utilize interaction recordings. Appropriated Knowledge assets such as citations are often treated as a special form of document-related sources and are, therefore, not detailed in the following, while some indicative examples for the others will be given.

3.3.2.1.1 Communication-Related Sources for Expertise Determination

Many approaches for expertise determination and later expert recommendation rely on e-mails as primary source of indications for expertise (Heeren, Sihn 2001; Krulwich et al. 1996; Oldigs-Kerber et al. 2003). The available systems operate in similar ways. Figure 3-8 shows a typical flow of steps.

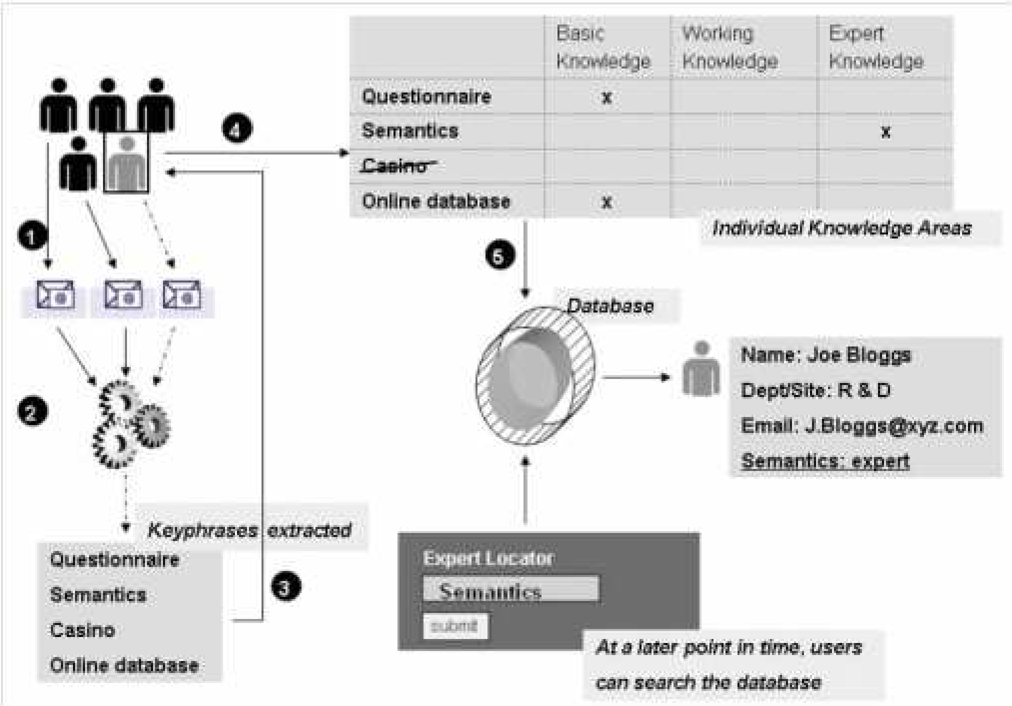


Figure 3-8 ***General idea of expertise determination using email messages***
Source: (Tedmori et al. 2006)

Whenever a users writes an e-mail it is passed through a system that extracts evidence for expertise. This is often done by extracting key phrases. There are different approaches how to find the relevant key phrases, such as the one described in (Tedmori et al. 2006; Campbell et al. 2003) or (Balog, de Rijke 2006). They differ in the algorithms they use for information retrieval or text retrieval or the parameterization of them. In a third step, the key phrases are sent to the original user and his expert profile. In step four, the user has the possibility to manually exclude them if he deems them inappropriate or alter the level of expertise with respect to those keywords. Step three and four are not present in all kinds of systems that rely on e-mail messages for expertise determination. In the last step, the information about the expertise and the person who holds it is then stored in a database – a system similar to a YPS – that can be queried when searching for an expert.

Using e-mails as source of evidence for expertise implies one assumption, which is also the same for other communication-related sources. One assumes that the knowledge that is *relevant* is communicated through the selected communication channel, i.e. in this case e-mail. Depending on the company size and its typical way of doing business, this might be justified. If, for example, a large company often relies on virtual teams and it is the company policy to use e-mails as principal means of communication, using just this media as source of expertise evidence is logical. Independent of whether or not one thinks that the textual contents sent via e-mails are valid and reliable indicators of someone's expertise, using e-mail messages as source of evidence has one distinct disadvantage. The extracted expertise is either not at all or only to a very limited degree *set into the usage context* of the e-mail sender. For example, an expertise relevant to a company might be "Programming Java". So, if someone sends and receives a lot of e-mails about programming Java he might be considered expert. For an expertise seeking individual the contextual factors however are very relevant. One individual that is assumed to be an expert in "Programming Java" due to his frequent use of the keywords might work in the HCM department and might try to staff projects with people who know how to program Java while another expert might be a programmer who often writes to his colleagues about issues faced when coding algorithms in the programming language Java. It now depends on the need of the expert seeking individual. If he needs someone who knows how to hire Java programmers the first expert would be appropriate but if he has a specific question related to Java programming in the sense of how to do programming in Java the second one would be more appropriate. This *lack of context* suggests to not use communication-based sources when designing this thesis' approach. However, the results of analyzing ERS that rely on communication-based source informs the later design of this thesis approach by emphasizing importance of context.

3.3.2.1.2 Document-Related Source for Expertise Determination

When using documents as a source for expertise determination, it is the assumption that the co-occurrence of relevant keywords and names of individuals are good indications of who knows what. Documents can be stored in databases, but also websites (Becerra-Fernandez 2000, 2001) can be seen as documents and can serve the purpose of feeding an ERS with necessary hints about expertise of users. Most of the first ERS relied on this source of information for populating their expertise directories. An early example can be found in Referral Web (Kautz et al. 1997) that analyzed the co-occurrence of names in documents. The names were arranged in a network based on a topic, which enabled the user to quickly find an expert for a given topic and indicated a way how to contact this individual through the common links in the network. In this way, Referral Web also illustrates a link to the related field of (social) network analysis.

The implicit assumption when using document-related sources for expertise determination is that the knowledge bearers *explicate their knowledge* or indicators thereof. While applicable in certain company contexts there is a problem with activities and the knowledge necessary for them that typically do not result in the creation of documents. For example, in a large international company an individual may be very experienced in trans-national accounting. The typical tasks in accounting however do not lead to documents where the expert would elaborate on his doings. Although he might do so via other means, the actual activity results in a document that typically does not contain his name and hence cannot be used as source for evidence. Owing to the fact that there are activities that do not necessarily result in documents that could be used for expertise determination, this thesis' approach will rely on a different source of evidence.

3.3.2.1.3 Interaction Recordings as Sources for Expertise Determination

Interaction recordings track the usage of software to deduce expertise from this interaction with the software. This kind of source for expertise determination is often used in software development. Here, expert mediation approaches may exploit information contained in the logs of version archives, that record the changes to source code that have been made by a programmer. The changes made are likely to reflect expertise with the methods that have been changed and methods that have been used (Schuler, Zimmermann 2008; Ma et al. 2009). Other approaches try to utilize knowledge evidence by looking at very basic interactions of users with a computer system such as mouse clicks and key strokes to elicit a user's current activity. The activity can then be stored as an indication for some level of expertise for this activity. (Beham et al. 2010).

The very idea of using interactions with a system as indication for expertise with respect to certain parts of the system seems promising. This thesis does not limit its scope to the area of software programming, however. On the other hand, the very generic approach to first determine a user's current activity by observing his behavior on a very detailed interaction level and using machine learning algorithms afterwards faces its own complexities. Correct learning of behavior from interactions with high levels of noise is error-prone. This thesis' approach adopts the idea of using interactions as basis for expertise determination as doing something is certainly a good indication of expertise concerning the activity. It may even be a *richer indicator* than the more indirect traces left by an activity such as follow-on communication or documents created in the meantime. Interactions also offer another advantage that will be utilized. The observation of IS interactions allows the *elicitation of a user's context*, which can additionally aid in fostering the exchange of knowledge. This thesis' unit of analysis will be an activity in general and in this sense, the resulting approach is more general than the approaches focusing on software development alone. However, it focuses on activities that are somehow reflected or can be reconstructed from IS logs. So, the resulting approach is more specific than the approaches that first try to elicit which activity a user is currently involved in. This way, the resulting approach reaches a suitable balance in the treatment of source data between very specific data source and hence strong knowledge of context for elicitation and on the other hand very general data sources and hence the challenge to first reconstruct an activity properly.

3.3.2.2 Approaches in ERS

From a user perspective, most ERS work similar to a "typical" search engine for document retrieval. The user enters search terms but is not presented documents that match the search terms, but rather individuals that are related to the documents, that contain the search terms (Craswell et al. 2001). More recently, some systems have emerged that do not provide a text-based search interface, but that use known individuals as "query items" to find similar experts for example as surrogates for already occupied experts (Balog, de Rijke 2007). While the use case and the interface are somewhat different, the principle means are the same for these kinds of ERS. To be able to offer either facility, it is necessary to have an intelligent means to determine expertise and to recommend those individuals that presumably hold this expertise.

The expertise recommendation can be done in different ways. Relating to the general view on the inner workings of recommender systems (see 3.3.1), most ERS rely on content-based recommendations. However, in contrast to general recommender systems, the context is slightly different for ERS. For the one part, the *user universe* and the *item universe* are the same. Individuals (user universe) are expecting the recommendation of other individuals (item universe). In addition, typically ERS do not ask the users to rate other individuals to use this information for the retrieval of suitable candidates. The *utility function* is, in this case, replaced with the search terms that are then used to compile a relevance value using a third universe: the sources used for determining expertise such as documents that have been created by individuals. This relationship is depicted in Figure 3-9.

Nevertheless, there are many different ways how the elicitation of expertise can be done. Serdyukov (2009) gives a good overview of common approaches that is summarized in the following. Additionally however, there are also ERS that can be categorized as knowledge-based recommender systems. They rely on explicit models about the item universe and/or the user universe.

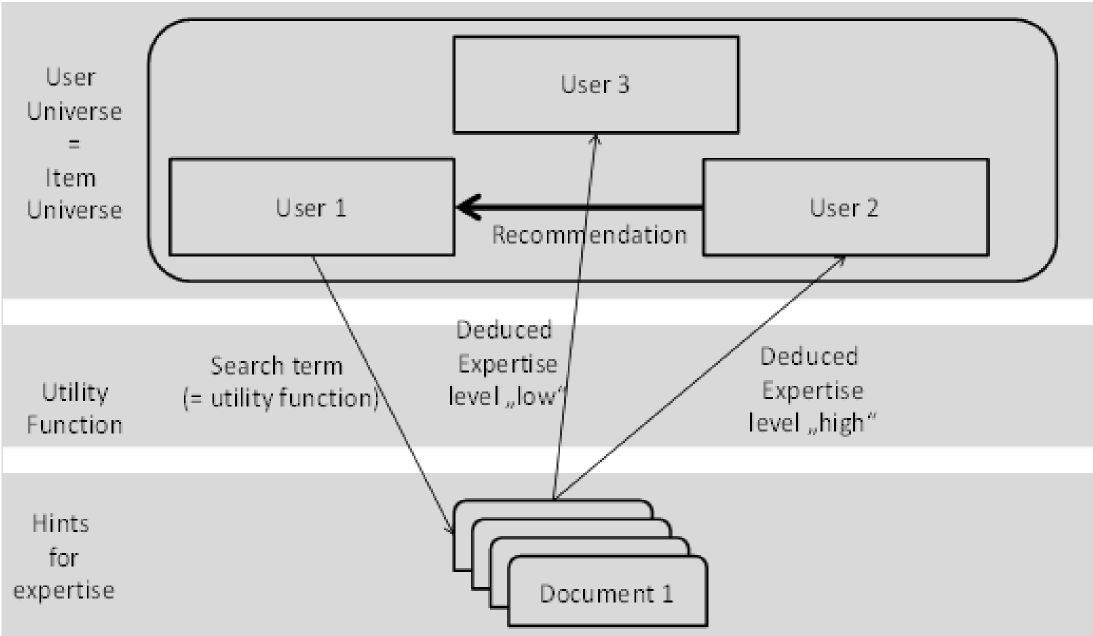


Figure 3-9 ***Expert recommendation as instantiation of content-based recommendation systems***
Source: Own Illustration

3.3.2.3 Profile-Based ERS

Early ERS used employees' profiles that contained textual information about their knowledge, skills, formal education, personal interests, external affiliations, and project membership as well as contact information. However, those were maintained manually. It was basically only a more sophisticated query front-end for YPS, that would relieve the user of the burden to search the profile pages. *Profile-based expert recommenders* can be seen as the first natural extension of this approach. The search for experts still indirectly relies on profiles of individuals. However, the individuals are relieved of the burden to keep their profiles up-to-date on their own, rectifying one of the supposedly biggest problems of YPS. The query for experts relies on explicit traces that are used to infer the knowledge of an individual as follows. Depending on the choice of sources that are deemed suitable in the organization's context, all traces of an individual, such as documents, files, emails and the like, that were either authored by an individual or that had a reference to the individual (by e.g. name or e-mail address) are merged into one "super-document". All super-documents that were created this way formed the basis for text retrieval algorithms with the knowledge seeker's request being the query terms, i.e. the super-documents were treated as one "normal" document each and a variety of algorithms commonly used in text retrieval were applied to rank the super-documents. The individual associated with the super-document that contains most of the requested search terms will be the top-ranked expert with respect to the query. The relevant research publications differ in the algorithms they adopt and which fine-tunings they apply to increase the precision of the retrieval.

The problems with this approach are twofold. First, predominantly relying on documents, e-mails and the like introduces the downsides as described before in sections 3.3.2.1.1 and 3.3.2.1.2. The other problem is rooted in the typical workings of information retrieval algorithms. The information contained in the occurrence of keywords is richer when there are fewer other words, i.e. it is the relative occurrence of keywords with respect to the remaining words that counts. As the construction of super-documents creates artificially long documents, the results of typical information retrieval algorithms are not optimal.

3.3.2.4 Document-Based ERS

Document-based expert finding strives to improve the profile-based approach. The idea to rank an individual higher with respect to the queried terms if those terms are represented more often in the super-document exhibited one downside as described before. The longer a document is, the lower the value of the information that one or more of the requested search terms are contained in it. This reciprocity in significance with respect to document length leads to the usage of document-based rankings. In this kind of approach, each document's significance with respect to the query terms is determined and the documents are brought into a ranked order. Subsequently, those individuals are presented as experts in a ranked list that have the most associations to the most relevant documents.

This approach reduces the problem of profile-based recommendation, in that it utilizes smaller units of analysis – the documents – to elicit expertise. However, the problems associated with the usage of documents, e-mails and the like as sources for expertise elicitation remain.

3.3.2.5 Window-Based ERS

Window-based expert finding approaches are a slight variation of the document-based approaches. Here, the co-occurrence of search term and person mentioned in a document is emphasized even more. The distance in words between relevant search terms and the name of an individual influences the probability that the individual is an expert, i.e. the more distant in a document a search term from a person, the less it is an indicator for expertise. Some other approaches also use strict limits, i.e. only those persons are counted that are within certain word bounds of the search terms.

This approach can be treated as a fine-tuning of the two approaches before. It strives to increase the value of information that is extracted from documents and the occurrence of keywords in them. From a conceptual point of view however, the approach is equivalent to the document-based expert finding approach.

3.3.2.6 Graph-Based ERS

The previous three approaches shared one commonality: the qualitative content of documents, e-mails and the like was used in one way or another to infer the probable expertise of individuals. Graph-based expert finding approaches go a different direction. They assume that someone is an expert if he has taken part in many interactions within a certain group. The interactions can rely on e-mail messages sent and received or more indirect forms of communication such as postings in online communities. It is less the number of keywords, their location or the pattern of occurrence that determines expertise, but the fact of having communicated about a certain topic. If, for example, an individual in an online community that focuses on a specific programming language has answered many questions of other members of this community, this person is assumed to be well-informed with respect to the overall communities knowledge base, i.e. the programming language. The communication graph can be exploited as source for potential expertise in different ways. Typical algorithms for this purpose use measures such as centrality in the graph, closeness or betweenness. However, the mere fact that someone has sent and/or received a lot of messages does not always indicate expertise. It especially does not necessarily correlate with a high degree of expertise.

By relying on past communication, graph-based expert finders have the merit of utilizing a social interaction as basis for expertise determination. In non-computerized settings the *word-of-mouth* about the expertise of an individual often serves the purpose of disseminating the know-who knowledge. This social interaction is reflected in electronic communications in the virtual world and communication graphs hence indicate who knows whom with respect to certain knowledge requests. However, graph-based approaches also face problems. Using the number of exchanged messages as criterion for expertise and hence for a decision on whom to contact, *implicitly excludes individuals* who might have high levels of expertise but have not (yet) exchanged their thoughts with others. On the other hand, individuals who are exchanging many messages concerning a certain topic are reinforced as being an expert by this approach. Their recommendation as an expert will increase the number of messages that the individual exchanges independent of the real level of expertise. This thesis' approach therefore does not rely on this source of expertise evidence. Rather it uses the event of doing an activity as evidence for expertise on this activity instead of the connected event of "talking" about an activity.

3.3.2.7 Model-Based ERS

In model-based expert finding systems, there exists a model that explicates what *areas of expertise* exist, which *levels of expertise* exist and possibly also relationships between areas of expertise that could be used for *inference mechanisms*. In the following, some examples are given to demonstrate the general idea.

Greer et al. (1998a; 1998b) report on two systems that explicitly rely on a model of tasks and the competencies that are associated with these tasks. The systems were used in a governmental unit as well as in a teaching setting at a university. The elements of their knowledge model can have various dependencies, i.e. one competence can be a prerequisite of another or one competence can be used to infer competence in another related competence. Building upon the knowledge model, in their system, a user’s expertise is modeled as an overlay to the knowledge model. Each individual has a certain level of proficiency in each competence item of the knowledge model, i.e. for each task. When a user of a system needs help executing a task, he can query the system with the task at hand as “search term”. The system then searches all users with higher proficiency levels than the user for this certain task. Greer et al. also describe some possible inference mechanisms. For example, if someone has high proficiency in an area of expertise, using their knowledge model’s dependencies, an inference algorithm could assume that the area of expertise that is a prerequisite for the first area of expertise also must be sufficiently well understood by the user. In their approach, the knowledge model for processes needs to be maintained manually once and adapted over time. They mentioned that the proper implementation of a knowledge model for one business process took some weeks and assumed that it will take approximately the same time for other business processes in their application domain. The competence overlay for each user is maintained by the individual user. Greer et al. reason that the individual is motivated to keep his competence profile up-to-date, so that he will only receive requests that he is able to help with. In sum, one can conclude that there is substantial effort for the initialization and maintenance of central knowledge models as well as for the individual competence overlay model. A problem that was shown to limit the success of YPS (see 3.2.3) and that similarly is a downside for this kind of ERS:

Beham et al. (2010) proposed a related approach with their system APODSLE. Their system’s functionality is based on a domain model. The domain model is an ontology that represents tasks within the respective domain, topics in that domain, skills that are required for performing the tasks, and a relation between tasks and skills that indicates which skills are necessary for which task. (Beham et al. 2010, p. 2786). While the modeling of tasks and necessary competencies exhibits some similarity to Greer et al.’s approach, the maintenance of individual expertise profiles is done differently. It is automated to a large degree, i.e. the level of expertise is not modeled manually, but inferred from the interactions of the user. For this purpose, the authors propose to use what they call *Knowledge Indicating Events*. An example for a knowledge indicating event could be, reading a document that is related to a certain area of knowledge, or helping someone with respect to a certain knowledge topic or previously having executed a task successfully. Additionally, the APODSLE system builds upon another assumption for suitable mediation of experts. To recommend knowledgeable people who can help in a specific situation, the system should take into account the current goal of a worker (e.g., a task and the demands that the task puts on the worker's skills) as well as the skills that a worker already has available. (Beham et al. 2010, p. 2784). At the same time, another design consideration is imparted in APODSLE. The system should be as unobtrusive as possible, determining the task context of the user automatically. This is achieved by applying machine learning algorithms that track the user's low-level interactions with the system, i.e. it monitors mouse clicks, open documents and web-browser interaction. During initial use, the user will be asked to indicate which task he is currently involved in and the system associates the low-level events with this task. After some time, the system learns which low-level events correspond to which task and is henceforth able to derive the current task context automatically (Lokaiczyk et al. 2007). In sum, one can conclude, that building a domain model requires some effort, as does the potential need to adapt it to future needs. The maintenance of individual user profiles on the other hand is automated as is the context determination after some initial training phase. In this way, the APODSLE approach promises to be a suitable way to reduce effort while still relying on model-based recommendation. However, the reduction in effort comes with the trade-off of potential misinterpretations in the context elicitation due to the machine learning model's flaws and potential misinterpretations of competencies when deriving them from knowledge indicating events alone.

Reich (2008) proposes another approach that relies on explicit models. In his dissertation he investigates potentials to support the execution of knowledge intensive processes, one aspect being the mediation of experts. He suggests relying on explicit models of the relevant knowledge intensive processes that contain steps in a process along with necessary competencies. In addition, he suggests relying on explicit models of an individual's competencies. One central design aspect of his approach is limiting the effort for initial setup of a framework as well as limiting the effort during productive use. For ease of initialization, he proposes to design the data structures of the process model and of the competence models in a way that is compatible with standards that companies currently use so that an easy import of existing data would be possible. For the competency profile, he reasons that semantic maps such as mind maps are an adequate structure and indicates that those should be importable easily. Another design aspect in his approach is utilizing the *wisdom of the crowds* also called *crowdsourcing* to keep process models, as well as the competence profiles of individuals up-to-date and correct. Users can grant editing rights to peers so that they may add or remove competencies or sub-competencies. Reich reasons that the individual is motivated to allow his peers maintain parts of the competence profile because it gives another dimension on the own competencies and thus shows the user things he might not be aware of himself. On the other hand, he reasons that peers are willing to maintain the profiles of the individual as a precise indication of the competencies of each individual will benefit the distribution of work within teams, which ultimately leads to a better team performance. Granting editing rights to more than one person promises to capture the reality more than relying on only one singular opinion. A similar mechanism is assumed to work for the collective maintenance of models for knowledge intensive processes. Here, Reich proposes to use tags to reflect that every individual might have slightly different views on the same process but indicates measures to aggregate similar tags of different individuals which ultimately also leads to an increase in the expressiveness of the model for the knowledge intensive process. In sum, one can conclude that Reich's approach needs some effort for initial modeling, whereby the extent of the effort depends on the availability of previously created process models in the company. The further maintenance of the process models once installed and the maintenance of the individual competence profiles causes ongoing effort. Reich's approach tries to limit the effort for the individual by distributing the cognitive effort onto the peers. However, since the individual still has to verify the collective edits in his profile and since the approach assumes that everyone also takes responsibility for the maintenance of others' profiles, it is unclear how much the effort will be reduced.

After having analyzed different approaches for model-based recommendation one can conclude the following. The initial setup of a model for tasks, i.e. a process model, the setup of a model for competencies and a possible hierarchy or similar relationship among those competencies as well as the setup of a competence profile for each individual is a *resource-intensive* task. There are approaches to limit the effort like Reich's idea to import already existing data or the APODSLE concept to derive competencies by relying on knowledge indicating events. However, considering Reich's approach, it is not always the case that companies have process models and it is also not clear if the models are sufficient as individuals might diverge from the process model (van der Aalst et al. 2005). The APODSLE approach cannot reduce the effort for the process model or competence model respectively but only for the individual competence profile. With respect to extended maintenance of the models, also *additional effort* is introduced. If there are only very infrequent changes to competencies or process models this might not pose a challenge. However, the more knowledge-based and not routine the work of an individual, the more likely it is that his competencies will evolve rather quickly and that the way of doing things in the company, i.e. the processes are adapted. Two ways of decreasing the inherent effort have been introduced: THE APODSLE system introduces some fuzziness by relying on machine learning algorithms while Reich suggests to distribute the effort among the team members. Although there are some suggestions how to reduce the effort, the principle challenge of coping with additional effort remains. This thesis' approach refrains from using explicit models of knowledge. It deduces likely expertise from the execution of activities, but does not use a formal model that indicates which activities need which levels of expertise. When not searching for knowledge carriers by entering keywords but by building upon the current context, an explication of expertise is not necessary. The current context replaces the formalization of expertise by assuming that an activity always requires the same kind of expertise and hence, when the context is similar to the context of someone who has executed an activity already, this is a valid proxy.

3.3.3 Advantages

ERS promise to offer a great potential for supporting the socialization process. They allow finding experts using interfaces that are similar to well-known search engines. In contrast to YPS, the data acquisition upon which expertise requests can be posted is automated. This relieves employees of the burden to maintain their profile pages manually, which often led to knowledge profiles, that were not up-to-date and consequently to the failure of YPS. Also, the setup of models of (hierarchies of) knowledge items and expertise is not necessary when building upon expert recommender. The source of expertise evidence itself will determine which keywords are relevant, or, if the company indicates relevant keywords for indexing, the input sources give good indications what are suitable keywords and hence reduce effort. An exception to this are model-based ERS that do use explicit models of knowledge and expertise as well as possibly models for work contexts and relationships between elements of knowledge or expertise. When again relying on models, the goal is to increase the accuracy, with the downside of having to maintain some elements of the system manually again.

The possibility to use arbitrary keywords for searching experts is also flexible in maintenance. Changes in the shared language or company business will be reflected in documents and other sources and hence the basis for searching experts evolves with these adaptations automatically. Additionally, ERS offer a variety of possibilities to tune the concept to the needs of the specific setting. ERS can operate on different sources for expertise determination. This makes them flexible for adapting the recommendations on the items that best reflect expertise in a given setting. Also the use of different approaches, such as relying on models or not, using content features of documents in various ways or relying on structures of social communication increases the flexibility and possible benefit of ERS.

Another advantage, especially in comparison to YPS is the fact that the willingness of experts to disclose their levels of expertise is not directly necessary. When relying on documents that are created during everyday work, or communicative acts like e-mail messaging that is also normal work behavior, ERS are non-obtrusive.

3.3.4 Disadvantages

While promising to offer valuable support for expert mediation, there are also challenges for ERS. The following five challenges are most problematic.

First, in many approaches, the expertise is elicited document-based, i.e. using documents that are in some digitally available format e.g. websites or files, or the expertise is elicited using communicative interactions such as e-mail messages or postings in a forum. There is an inherent problem in this case. It is assumed that the connection between digital documents or communicative interaction and individuals is a *reliable source* for inferring expertise. The argument why this is a challenge may go two ways. Firstly, it might be the case that someone authored many documents, but the content was not of his origin. The higher someone is within a company's hierarchy the more often his subordinates will have created the actual content. Secondly, it might be the case that someone is very knowledgeable in a certain topic but did not leave many digital traces, i.e. documents or messages that could be used for inferring his expertise.

A second problem lies within the definition of "being an expert" in the context of ERS. The ERS can only rely on documents that are available to it to infer the relative expertise of employees. Absolute expertise is on the other hand not inferable. While the definition of (absolute) expertise on a more global level is problematic in itself, the reliance on only relative expertise still is a challenge. Under the given context, ERS can only yield a best possible fit not necessarily a good fit. If no one in the company has knowledge about a certain topic, the one with very little knowledge would be considered expert – among the blind the one-eyed man is king. Yet, there is also a challenge in the contrary direction. ERS rank the relative expertise suggesting the most expert individuals that are available. However, not always is it necessary to use the most expert individual nor is it always beneficial. If someone has a moderately complex question it does not need a company's most brilliant expert in this area, another coworker might be suitable in terms of level of expertise. Framed differently, ERS typically do not address the efficient use of the "resource expert" but normally always indicate the most proficient individuals. Additionally, the most knowledgeable individual might not be a good selection after all. Independent of the cost of consulting a top-expert, this top-expert might not be able to help equally well as a less experienced individual that knows how to convey the necessary knowledge by having a higher overlap in shared mental models. An example from private life shall illustrate this. If a family is looking for a private tutor to help their child learn physics, they often ask a pupil from a somewhat higher class or maybe a student from university. This makes sense as pupils or students are in *similar situations* in terms of learning the concepts of physics. Typically, no one would think of asking leading scientists like a Nobel prize winner to tutor school kids. However, this is exactly what typical ERS would do. Very likely the knowledge transfer would be hindered tremendously by the lack of shared mental models and shared language rendering the "expert" not an appropriate help for the specific need. In cognitive psychology, this phenomenon is referred to as *knowledge de-encapsulation* and de-chunking: experts store their knowledge in higher aggregated chunks, i.e. they encapsulate knowledge that has become trivial for them in chunks. To be able to help, they have to unpack these chunks which is a high mental effort as they basically must undo their previous learning to understand the requester's problem context (Bromme et al. 2004).

A third challenge lies in the general idea of bringing together knowledge seeker and knowledge bearer for a face-to-face meeting or at the least for a *synchronous meeting* via e.g. teleconferencing software or phone calls. Most solutions and research efforts in the area of expert recommendation focus on expert identification and less so on expert selection with its two subtasks expert localization and meeting scheduling (Serdyukov et al. 2008) and hence do not address the question of how to actually connect the knowledge seeker and the knowledge bearer. Serdyukov (2008) describes some possibilities for this purpose. However, the assumption, that synchronous interaction is necessary, which is the original idea of the socialization process, is not always true – sometimes asynchronous interactions suffice, e.g. in case when someone asks an expert for yet undocumented knowledge. Request and answer do not necessarily have to be synchronous for knowledge exchange to happen in this case. ERS however focus on mediating synchronous communication. Additionally, expert recommender typically offer no mechanism to overcome *social inhibitors* to addressing an expert, such as *fear of losing one's face* for asking someone else, or on the other end, *motivation to answer requests*.

Fourth, ERS are typically systems on their own. They are not integrated into the operational systems that individuals use for their actual work. On the one part, this means there is a change in media or system used, which might be an additional burden for the user. On the other part, this means that the individual must know what he needs to know, i.e. he must be *able to articulate* his need for an expert by realizing which competence or skill he currently lacks. While this is more problematic with YPS, where the user has to stick to a given set of competencies, the challenge remains in ERS to a somewhat smaller degree. Here, the user can iterate through documents to narrow down suitable keywords. However, without integration into an operational system, there is no way to support this mental process as the context of the individual cannot be elicited. Exceptions to this can be found in some prototypes (Beham et al. 2010) that try to elicit the user's context, but also in these systems, the implementation is an additional tool that is not within the operational system but that only observes the operational system.

Fifth, as was true for YPS, the goal of ERS is the facilitation of the socialization process. However the focus is still on *one-to-one interaction*, i.e. one individual uses the ERS to find possibly many suitable candidates for querying but usually only queries one. This leads to a relatively unmanaged network of knowledge flows between individuals (Oldigs-Kerber 2007, p. 67), where the fruits of the interaction are neither stored electronically in some form, nor in the heads of others but the two interaction partners.

The problems of expert recommender are summarized in Table 3-3.

Id	Description
Problem 1	Relationship of keywords and person not necessarily authoritative
Problem 2	Definition of expertness not with respect to context and current need but global
Problem 3	Expert selection often not supported; no concept for supporting asynchronous interactions
Problem 4	Non-integrated which necessitates change of media; very seldom consideration of user context
Problem 5	Creates one-on-one interactions, that are not well manageable

Table 3-3 ***Summary of problems that expert recommender face***
Source: Own Illustration

3.3.5 Summary

Similar to YPS, ERS try to support the transfer of knowledge through mediating the socialization process. To find a suitable expert for this purpose and also taking into account the large number of possible individuals, the system creates recommendations for possible experts. This information filtering support is part of active research in the more general class of recommender systems. Recommender systems apply different approaches such as recommending *similar items* with respect to those that are known to the user and that had high utility for him. Other recommender systems find *similar users* and recommend what those users liked. There are also *hybrid combinations* of these two approaches. Finally, there are *knowledge-based recommender* that try to increase the value for their users by taking information from a previously created knowledge base that contains information about the item universe, the user universe and links between the two.

Users that utilize ERS search for the *necessary competencies* similar to searching for documents, i.e. by entering keywords. The necessary data to search in, i.e. expertise of individuals, is gathered automatically to relieve individuals from the effort of maintaining profiles. This was a challenge that is assumed to be a large hindrance for the acceptance of YPS. To automatically gather the data, ERS can rely on different kinds of source data. *Document-based sources* subsume files, websites and the like, while *communication-based* sources record communicative interactions and their content such as e-mail messages that have been exchanged or messages in a forum. A last source of possible expertise evidence can be found in *interaction recordings* of programs. ERS can either rely on *explicated knowledge*, independent of its source being documents or communications, or they can rely on connections such as *graphs of communications*. Finally, expert recommender can be *model-based*, i.e. they can utilize models of knowledge items, users or contextual elements to increase their effectiveness.

Analyzing the areas where ERS face challenges one can conclude some learning for creating a more suitable KM approach. First, the relationship of expertise evidence and actual expertise needs to be more authoritative. While in some settings, the co-occurrence of keyword and person in an e-mail or a document may very well indicate some level of expertise, this is not always the case and it is problematic, when e-mails and documents are not the natural output of the knowledge work. It would be beneficial to have sources of expertise evidence that are *inherently bound to the activity* of the users and that reflect actual doing rather than some indirect result such as documents. Second, the matching of who is suitable as an expert needs to be more dimensional taking into account not only the determined absolute expertness but also the *contextual relationship* between knowledge bearer and knowledge seeker. Third, contacting experts should be made easier. Not only synchronous communication needs to be supported but often *asynchronous* is sufficient but less obtrusive for the expert and thus it increases the likelihood of acceptance on the expert's side. The selection of a suitable expert should furthermore be supported by taking into account the *context of a user*. Fourth, the individual should be supported in the search by allowing the context to influence the search and this way to *reduce the mental effort*. Change of media should be avoided. Fifth, there should be means to generate a learning for the company as a whole in some way additional to the learning that takes place between the two interaction partners.

The conclusions drawn are summarized in Table 3-4

Id	Description
Conclusion 1	Expertness should be determined building upon more direct evidence such as users' activities
Conclusion 2	Absolute level of expertness should not be used alone to recommend "experts"; shared mental models should be considered also
Conclusion 3	Means for connecting knowledge seeker and knowledge bearer should make use of both user's context taking into account social aspects
Conclusion 4	The current context of an individual should support the search process; recommendations should be possible from within the original work context
Conclusion 5	Means to increase the benefit from the interaction to more than just the two partners are desirable

Table 3-4 ***Summary of conclusions from analyzing expert recommender***
Source: Own Illustration

3.4 Knowledge Networks

As mentioned at some times throughout the thesis, most of the knowledge in a company is not available in digital or otherwise explicated format but is only "stored" in the heads of its employees (Krcmar 2010, p. 663). In addition, earlier noted, next to supporting *knowledge flows*, supporting the *creation of knowledge* is a goal in KM (Probst et al. 2000). With respect to tacit knowledge, both these goals can be obtained by relying on socialization processes (Nonaka, Takeuchi 1997). However, while the two approaches described before – YPS and ERS – put their emphasis on facilitating one-on-one interactions, the idea of knowledge networks is to focus on bringing together a larger set of individuals that interact with each other. Through their interactions, knowledge is transferred from one individual to one or more others. Additionally, new knowledge is created typically not by individuals but in discussions and interactions of groups of individuals. Therefore, the interaction allows increasing the collective knowledge – for each participant and the organization. It is also a very natural way of fulfilling one's task to rely on personal networks of co-workers for additional support. In fact, after investigating the work behavior in some studies, the awareness surfaced that work-related information is to a large extend passed on informally which was another supporting fact for nurturing communities for knowledge exchange in a more systematic way (Petter et al. 2007, p. 2).

3.4.1 Clarification of Concepts

While the principle idea of individuals gathering for mutual knowledge exchange is ages old (Wenger, Snyder 2000), the question how to *setup*, *manage* and *support* those sets of individuals has been in the focus of (knowledge) management and organizational learning research in recent years. A number of concepts have emerged under sometimes similar names but with different assumptions underpinning the approaches and, on the other hand, with different names but similar interpretation. Some of the terms that are related to sets of individuals and their interaction are:

- Virtual Communities
- Communities of Practice (CoP)
- Virtual Community of Practice (VCoP)
- Online Community of Practice (OCoP)
- Computer-mediated Community of Practice
- Electronic Community of Practice (ECoP)
- Distributed Community of Practice
- Knowledge Communities
- Knowledge Networks
- Networks of Practice
- Communities of Interest
- Communities for Innovation

The use of this many terms without clear distinction between their characteristics is confusing and problematic for the field of research (Teigland 2003; Teigland, Wasko 2006). Therefore, there have been attempts to structure the different kinds of social fabrics that can emerge.

Noting that there are many terms referring to similar phenomena, Leimeister (2004) tried to define and classify virtual social interaction. In his dissertation thesis, Leimeister researched the design and operation of virtual communities. He analyzed thirteen publications about virtual communities to find and synthesize suitable dimensions for categorizing them. Finding that the dimensions and categories are not always mutual exclusive he states his own working definition of virtual communities. A virtual community is a sub-form of *social groups* that subsumes individuals with a *desire for information or interaction* or the desire to assume a certain role in a group. The connecting element between the individuals is a *common idea or goal* (which includes task, problem or interest). The social construct has *implicit or explicit rules of conduct* and the interaction is mediated by a *technical subsystem* that is capable of establishing *trust* and a *feeling of identity* for its users, without them being physically present. It is a socio-technical system (Leimeister 2004, p. 37). This definition also fits for social groups that focus on knowledge exchange as Leimeister notes.

Teigland (2003) sees the term *network of practice* as the most general term for groups that form around knowledge needs and proposes to structure the possible instantiations according to two dimensions – the primary means of communication and the nature of the network of practice. She identified five different kinds of social fabrics that she classified along the two dimensions, as depicted in Figure 3-10.

While Teigland sees communities of practice as one special instantiation of networks of practice, Archer (Archer 2006) on the other hand sees the term community of practice as the domain of interest and suggests to distinguish communities of practice further according to the structure of the organization that they serve. He proposes to distinguish four forms: Internal Community of Practice, Communities of practice in network organizations, formal networks of practice and self-organizing networks of practice. The differentiation is similar to Teigland's "nature of network of practice" dimension. Finally, Koch et al. (Koch et al. 2007) argue that there is no clear distinction between (virtual) communities of practice and knowledge networks but rather they fluidly converge into one another.

In the further discussion, the attempts for categorization are synthesized with respect to this thesis' goal of fostering the knowledge flow between a knowledge seeker and a knowledge bearer in an organization. A *common interest* is relevant for a social interaction to happen. This was noted in all categorizations. Further one can note that the involvement can vary, depending on organizational aspects (Teigland's "nature of network of practice", Archer's "structure of the organization") and on the possible medias to interact (Teigland's "primary communication media"). In this thesis, the investigation will focus on three instantiations, first the "classical" community of practice, that includes face-to-face meeting, and builds upon strong social ties among its members, second distributed networks of individuals that meet every now and then in person but predominantly interact through other means, and third loose networks of individuals that exchange knowledge on a non-regular basis and most often never meet in person.

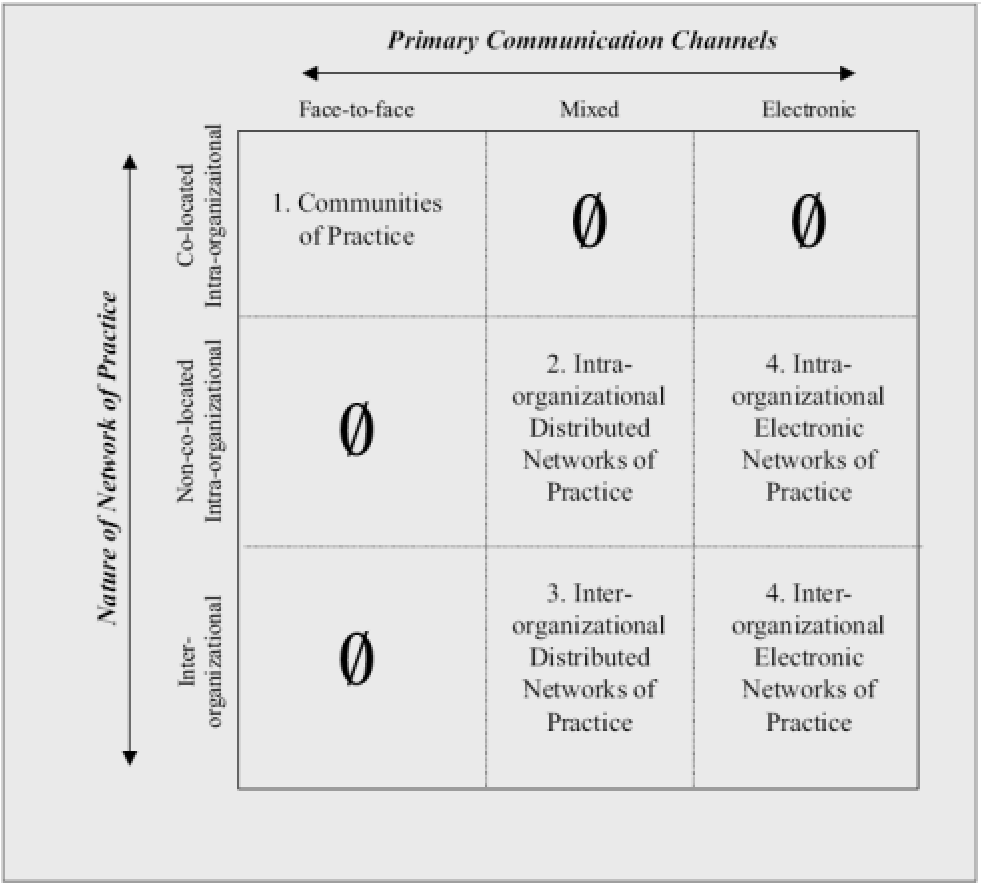


Figure 3-10 ***Matrix of Networks of Practice***
Source: ***(Teigland 2003, p. 24)***

3.4.2 Community of Practice

This section will shed light on a KM approach that has attracted much attention especially in larger corporations. First, this section strives to find a suitable definition of the term Community of Practice. Then, the aspects of Communities of Practice that can be influenced, i.e. the elements that can be managed are investigated to understand what their potential and disadvantages for KM are. Further, the constituting elements of a community are worked out to later contrast them with similar yet different forms of social interaction that are also part of KM research.

Communities of practice (CoP) is a term that started to attract attention beginning in the early 1990s (Lave, Wenger 1991; Brown, Duguid 1991; Wenger 1998; Wenger et al. 2002).

While the underlying concept is as old as social interaction among human beings, the special attention in terms of understanding the concept and thinking of how to support and capitalize those social interactions is new.

The fundamentals of research in the field of CoP can be found in Lave and Wenger's shift in analysis of how learning happens that led to the investigation of situated learning (Lave, Wenger 1991). Analyzing the learning that happens in apprenticeships, they realized that the apprentice learns not only from the master but also from other apprentices with only some more experience. Lave and Wenger state that all of them, the master, the apprentice and the more advanced apprentices, form a CoP – a social construct that is about what people do in their profession. The concept however not only holds for apprenticeships but for all learning situations that involve social interaction. Furthermore, Lave and Wenger postulate that learning is not simply transferring or "copying" knowledge from the more knowledgeable individual to some receiver, stating that „ […] learning as internalization is too easily construed as an unproblematic process of absorbing the given, as a matter of transmission and assimilation." (Lave, Wenger 1991, p. 47). Rather the context determines the learning and CoPs are the place – the situated context – for the social learning (Wenger 1998).

3.4.2.1 Definition of a CoP

In an attempt to more precisely define what a CoP is, Wenger et al. state that CoPs are " […] groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis […] they accumulate knowledge, they become informally bound by the value that they find in learning together" (Wenger et al. 2002, pp. 4–5). In the German-speaking community North et al. (2000) also researched the phenomenon of CoP but named it differently. They refer to knowledge communities, which they define as groups of persons that exist over an *extended period of time*, that have a *joint interest* on a specific topic and that want to *collectively create and exchange knowledge* about it. The *membership is voluntary* and *personal*. Knowledge communities are arranged around specific content (North et al. 2000).

CoPs are informal organizational structures in which membership is not mandated but is voluntary. They are additional to typical structures in an organization and together with them form the *double-knit knowledge organization* (Wenger et al. 2002). Figure 3-11 illustrates the interaction between CoPs and other forms of integration in an organization. Individuals work in groups that are dedicated to operational tasks, such as workgroups or teams. The knowledge they acquire there can be brought into a CoP to share it with other like-minded individuals either by discussing it directly or by documenting it to create the possibility to discuss. In both cases, the CoP can validate what the individual has contributed. This in turn leads to quality assurance of the knowledge that has built up in the CoP. Furthermore, the ability to solve problems among all CoP members is increased which results in better performance in their operational groups. This is where the cycle reiterates. The new skills enable individuals to perform their work better, possibly adapting it to the practices they have learned through the CoP until they find new challenges that they master by relying on their CoP's support.

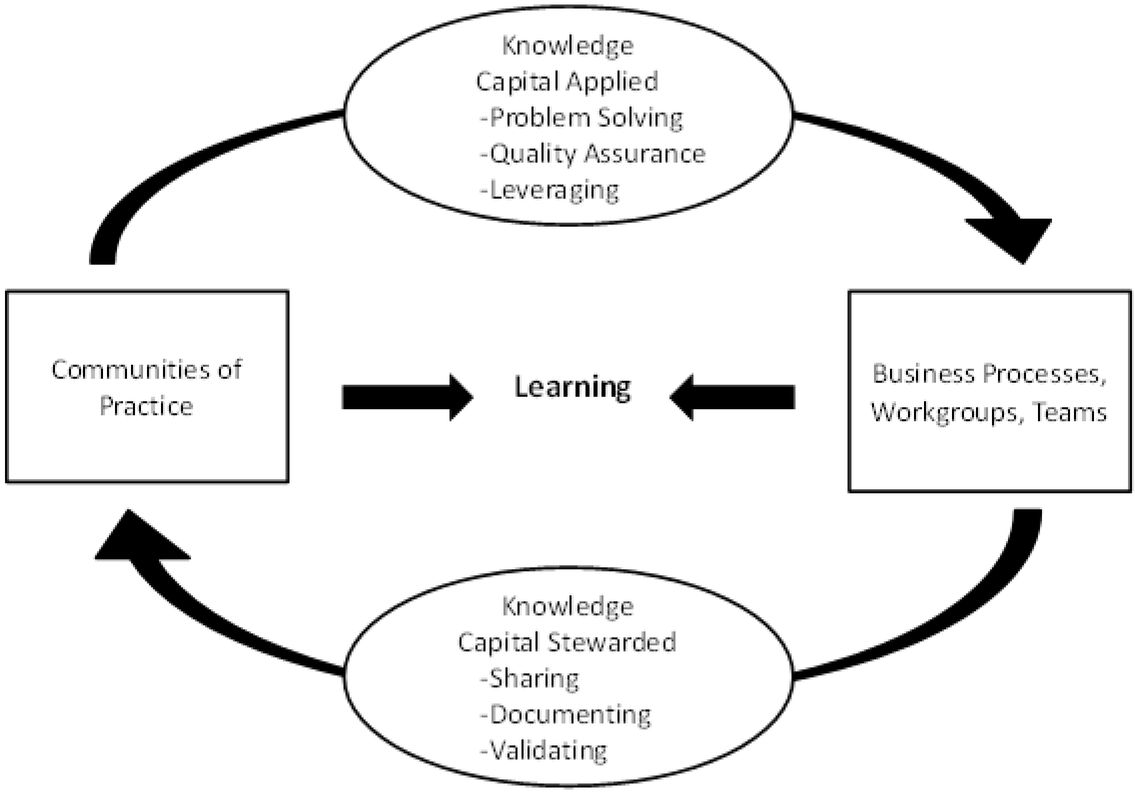


Figure 3-11 ***The double-knit knowledge organization***
Source: ***(Wenger et al. 2002, p. 19)***

Wenger further indicates that "A CoP defines itself along three dimensions: its joint enterprise as understood and continually renegotiated by its members, the relationships of mutual engagement that bind members together into a social entity, the shared repertoire of communal resources (routines, sensibilities, artifacts, vocabulary, styles, etc.) that members have developed over time." (Wenger 1998, p. 2). Later he relaxes this definition and merely states that a CoP is determined by three interdependent structural components: *domain*, *community* and *practice* (Wenger et al. 2002, pp. 27–38). The domain is a set of topics, focal points, problems and open issues (Bettoni et al. 2004, p. 320) that builds the common ground for the CoP's members and allows them to identify with the CoP. It informs its members what can and should be brought up in the community and which state of maturity of contributions is expected. Examples of topics can be specific technologies, processes, methods, products or personal experiences (North et al. 2000). The community comprises the CoP's members, their relationships and interactions. It is *the social fabric of learning* (Wenger et al. 2002, p. 28) that is based on mutual respect and trust. In accordance to the model of *situated learning* (Lave, Wenger 1991) a CoP's community is where learning is situated, i.e. it is a matter of belonging as well as an intellectual property (Wenger et al. 2002, p. 28). The practice comprises artifacts that, socially defined, indicate the ways how to do things in a specific domain. It includes both tacit as well as explicit knowledge such as a set of common approaches, shared standards, reference models, lessons learned, heuristics or instruments (Wenger et al. 2002, p. 38).

The three structural elements can also be used to differentiate CoPs from other organizational groups as illustrated in Table 3-5. Formal work groups, which are part of the right hand side of Figure 3-11 are concerned with a specific task, which could also hold for a CoP. However, membership within a work group is not voluntary and it is not the passion for the specific task that indicates membership but the job description. Project teams typically focus on a specific task, but they vanish once a project is over, while CoPs tend to live as long as they are able to contribute value to their members, which is typically much longer. Also in project teams, the members are typically assigned and do not join (purely) voluntarily. Yet, the mere fact that someone joins a group voluntarily does not alone constitute a CoP, nor does the additional fact of sharing a passion about a specific topic. Informal networks fulfill both those requirements, but they are different to CoPs in that they do not focus on progressing the knowledge and expertise of its members but rather function as conduits of information.

	Purpose	**Members**	**Coherence through**	**Lifetime**
CoP	To develop members' capabilities; to build and exchange knowledge	Members who select themselves	Passion, commitment, and identification with the group's expertise	As long as there is interest in maintaining the group
Formal work group	To deliver a product or service	Everyone who reports to the group's manager	Job requirements and common goals	Until the next reorganization
Project team	To accomplish a specified task	Employees assigned by senior management	The project's milestones and goals	Until the project has been completed
Informal network	To collect and pass on business information	Friends and business acquaintances	Mutual needs	As long as people have a reason to connect

Table 3-5 ***Differentiation of CoP towards other organizational groups***
Source: adapted from ***(Wenger, Snyder 2000, p. 142)***

3.4.2.2 Managing a CoP

CoPs are self-establishing, self-maintaining and self-managing by definition. Hence to utilize the concept in a direct way for KM it is necessary to understand how they can be influenced at all, i.e. how they can be managed. Not surprising, the question of how to establish and maintain CoPs has been the focus of much debate. North et al. (2000) propose that a suitable context can be created that increases the likelihood of the establishment of a CoP. They arrange the means of interaction according to four dimensions: *persons,* that are members of the CoP, the *interactions* between those individuals, the resulting *knowledge transformation,* and the *organizational embeddedness,* stating that some, like membership criteria can be influenced more directly than others like motivation of participants. Wenger et al. (2002) also indicate that CoPs cannot be created but their creation can be fostered and their growth nurtured. As they see CoPs as living elements, they indicate seven guidelines that should support the CoPs' development. In the following North et al.'s four dimensions and their aspects to nurture CoPs are described. For each dimension, the analogies to Wenger et al.'s seven guidelines for nurturing CoPs are outlined. The relationships between both author groups' elements is also depicted in Figure 3-12.

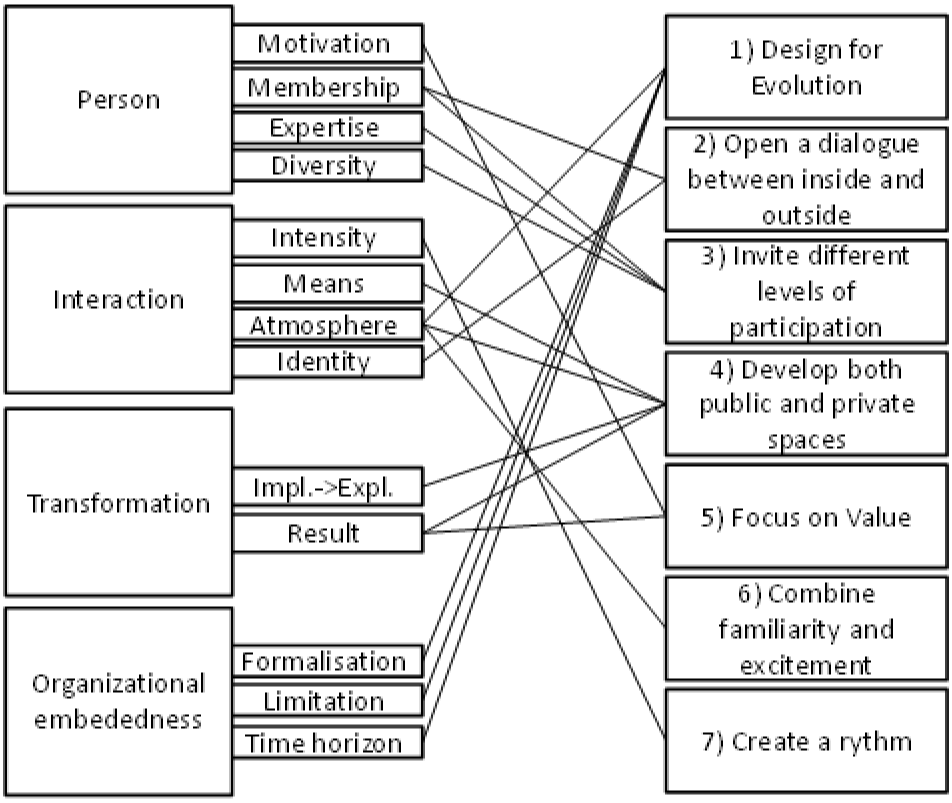


Figure 3-12 ***Relation between North et al.'s dimensions for CoP management and Wenger et al.'s guidelines for CoP management***
Source: Own illustration in accordance with ***(Wenger et al. 2002, pp. 51–63)*** and ***(North et al. 2000)***

Within North et al.'s context dimension *person*, the aspects *motivation*, *membership*, *level of expertise* and *level of diversity* can be influenced. The motivation can be influenced indirectly by e.g. asking for challenging goals for the company or work group but this measure is typically only short-lived (North et al. 2000). Goal setting is more productive when done by work groups themselves and the highest motivation is to be expected when membership in a CoP is positive for the individual work situation. This is obviously in line with the general idea of Wenger et al. (1998) who highlight this interaction as the actual benefit of CoPs. Wenger et al.'s guideline 5 indicates that the value for the individuals should be focused. Membership is typically regulated by the CoP itself, which is the original intent. North et al. mention that mandated membership is possible but will likely create problems. The mandated persons might not be accepted by the group and seen as a foreign particle. If mandated, trust in the decision to include a new member on mandate needs to be obtained. Also, participation will vary in its degree. CoPs can grow to an extensive size which leads to different levels of participation. If CoPs grow too large a segmentation into sub-CoPs often happens and is beneficial to its members. Wenger et al. on the other hand state that interaction with externals help the community understand what it is about and who should be a member. In addition, they suggest to attract individuals with different willingness to contribute as not only the core team creates value for the company. This last statement is taken up in North et al.'s aspect of *level of expertise* and *diversity*. The level of expertise is another field for intervention according to North et al. that indicates if the members are familiar with a topic, know much about it or if they are experts. However, they do not state how to influence this criteria, i.e. how the selection of varying levels of expertise should be done. The same holds for the level of diversity. North et al. indicate that the mixture of different backgrounds influences the CoP as such but do not indicate any proposals for how to adapt the level of diversity for a CoP to be successful.

Within North et al.'s context dimension *interaction*, the aspects *intensity*, *means of communication*, *atmosphere*, and *identity* can be influenced. The intensity describes the frequency and duration of meetings or contacts of CoP members. Wenger et al. also note that intensity needs to be managed by referring to a suitable *rhythm of the CoP*. It determines how often face-to-face events should be scheduled and when regular interactions should take place The means of communication determines how the CoP members interact, e.g. by phone, video conference or e-mail. North et al. state that personally knowing someone strongly facilitates later communication via electronic media. Further, they note that the more implicit the knowledge that is core of the CoP the more there needs to be personal communication face-to-face. Wenger et al. emphasize this even more by stating that a major part of a CoP's value lies in the network of one-to-one connections that is built up, which is why they explicitly suggest to create communication means for the CoP's public and for one-on-one interactions.

The collaboration between individuals is determined by the atmosphere of the interactions. Trust and openness can however not be mandated, they must develop, which in turn can be supported by providing suitable conditions. Rules of conduct can be established and the place where interactions take place also increases or decreases the feeling of intimacy that determines the atmosphere. Wenger et al. formulate this by asking for familiarity to be fostered but also including excitement, i.e. new opinions. Also, the atmosphere can be more open when one-on-one interactions are fostered. Another aspect that is related to the discussion before is the identity of the CoP. It can be fostered by e.g. providing websites where the CoP can describe itself, or by designing and assigning a logo to the CoP. Over time, other characteristics like CoP-specific languages or rules will evolve that also help to differentiate the CoP from the rest of the organization and this way to give rise to its own identity. Wenger et al.'s guidelines adopt this with their suggestion to have an open dialogue between inside and outside of the CoP which helps to identify what constitutes members of the CoP.

Within North et al.'s context dimension *knowledge transformation*, the aspects *transformation of tacit to explicit*, and *result* can be influenced. With respect to the transformation processes, North et al. indicate that next to the transmission of knowledge within a CoP, tacit knowledge is created as well. However, the results are not always interpretable by outsiders of the CoP as the process of creating the knowledge artifact is implicit knowledge in itself. The results of knowledge transformation can be assessed using guiding questions, such as "what can I use in my daily work" on the individual level or "which parts of our work are most asked for" on the level of the CoP itself. This reflection of the results shows the integration of the CoP into the organization's operational units. However, North et al. do not indicate any means of influencing this dimension, they merely state that it is a result of the interactions. Wenger et al. indicate that the result will be influenced by trying to build CoPs according to fields that contribute value to the overall organization. The transformation of explicit to implicit knowledge is in their guidelines also supported by the one-on-one interactions that are facilitated by offering private interaction possibilities.

Within North et al.'s context dimension *organizational embeddedness,* the aspects *degree of formalization*, *limitation* and *time horizon* can be influenced. The formalization indicates which visibility a CoP has within the organization that it belongs to. North et al. adopt Wenger et al.'s five level categorization. In the first level, CoPs are invisible to the organization and maybe also to its members that are just implicitly "members". In the second level, called "bootlagged", the CoP is visible only to a small group in the respective field. In the third level, the CoP is legitimated by the organization and seen as a valuable entity. In the fourth level the CoP is regarded as strategic; it is respected as a central entity for the success of the company. In the final, fifth level CoPs are transformative; they poses the ability to redefine their surrounding organization and alter the direction the organization goes. In terms of the limitation of CoPs, North et al. indicate that CoPs can be within a department, cross-departmental or even cross-organizational. Also in terms of geographical aspects, its members can be in the same local setting, in the same region, country or they can be internationally dispersed. Further, CoPs can be distinguished from each other, since there will be more than only one CoP and normally there will be some overlap in topics and members. Finally, the time horizon of a CoP determines its contribution to the overall organization. The reason d'être of a CoP can vanish over time when the context of the organization changes. Wenger et al. address these aspects in their guideline to design a CoP for evolution. They reason that the context changes over time but so does the CoP itself.

3.4.2.3 Community Aspects of CoPs

In order to better understand the difference of CoPs to similar other social constructs, that exhibit less intense relations between their members, it is necessary to investigate what a community is. Wenger et al. (1998) as well as North et al. (2000) have used the term community to denote the social fabric that they describe. Wenger et al. also explicitly state the difference to other forms of social groups, nevertheless they do not indicate what a community is in their eyes. One interpretation of what a community is can be found in (Bender 1982, p. 7): "A community involves a limited number of people in a somewhat restricted social space or network held together by shared understandings and a sense of obligation. Relationships are close, often intimate, and usually face-to-face. Individuals are bound together by affective or emotional ties rather than by a perception of individual self-interest. There is a 'we-ness' in a community; one is a member." So, according to (Bender 1982) communities are *limited in size*. Wenger et al. indicated this as well but noted that there may also be large CoPs with hundreds or thousands of members. What is missing is a rough number of how large CoPs can be or should be to still function. Further, communities have a *shared understanding* and *sense of obligation*. In CoPs the shared understanding is built up throughout the existence of the CoP by *continuous renegotiation* of meaning, ideas and interpretations. The sense of obligation in CoPs is not equally elaborate for all individuals. Wenger et al. note that around 10 - 15 % of all members form the core, another 15 - 20 % form an active group around the core and the rest is less active or not active at all but still members of the CoP (Wenger et al. 2002, p. 56). The *closeness* and *intimacy* of relations is also stressed in Wenger et al.'s descriptions of CoPs. Regular meetings, typically face-to-face, are in their eyes indispensible for fostering a feeling of "we". Other definitions of community also stress that the relationships between the members are *affect-laden* and that all members are committed to *shared values and norms* as well as a (at least partially) *shared history* and a *shared identity* that is granted through the community (Etzioni 1996). It is the close interaction and the *personal acquaintance* that forms the basis for CoPs. The more the individuals interact and contribute to the CoP the more trust builds up among its members. Increased levels of trust foster the individuals' willingness to share knowledge with others in the CoP. In addition, one motivation for individuals within CoPs is fostering their relationships by contributing their knowledge (Davenport, Hall 2002, p. 192). Further, in line with the intimate relationships among a limited number of individuals, in CoPs *expectation of reciprocity* is strong, and individuals are aware of each other's actions, resulting in a relatively high degree of *social control* (Teigland 2003).

3.4.3 Virtual/Online/Computer-mediated/Electronic/Distributed Community of Practice

CoPs in their original definition require their members to learn from each other through frequent, intimate interactions which requires also frequent face-to-face meetings (Teigland 2003, p. 33). The concept of a community can however also be transferred in the online world, i.e. the interactions between the individuals can be based predominantly on non-face-to-face interactions such as e-mails or bulletin boards. However as Teigland and Wasko note (2006, p. 188), "[…] online social structures focused on knowledge exchange are commonly referred to by numerous names: virtual or electronic communities, communities of interest and online communities". The common theme is that they are mediated through some means, typically using *computer-mediated* or more colloquial *electronic* media and hence interaction is *virtual*, more colloquial it is done *online* connecting locally *distributed* individuals, which is why all of these terms can be seen as synonyms, although some authors connote sometimes slightly different meaning to each respective term. In this thesis, these terms will be used interchangeably.

As the members of such virtual CoP do not rely on face-to-face interactions as the primary means of interaction, it is relevant to inspect which aspects are affected by this change. Thiedeke has investigated which sociological changes are introduced once one relies on computer-mediated means of communication that he implies by referring to the word virtual. He identified the following four areas of difference (Thiedeke 2003, pp. 25–33):

- *Anonymity*: In purely virtual settings, it is not possible to identify the interaction partner.
- *Losing self-confining*: In purely virtual settings, it is possible to deviate from rules of conduct and to leave social roles as there are no means for sanctioning this behavior.
- *Interactivity*: Computer-mediated means of communication allow communicating with more than one individual.
- *Optionality*: Individuals can choose the topics, forms, and intensity of interactions.

Certainly, this change in sociological aspects will influence the behavior and interactions between individuals. The discussion will focus next on different instantiations of virtualized CoPs, with varying degrees of "virtualization". In the following the terminology of Teigland (2003) will be used to refer to these different instantiations. She refers to the social constructs as networks of practice rather than CoPs. This makes sense as section 3.4.2.3 showed that communities exhibit strong (emotional) ties between their members, which may not hold for more virtualized forms of interaction.

3.4.4 Distributed Networks of Practice

Distributed Networks of Practice comprise individuals "[…] who are dispersed across the organization yet who work on similar tasks using a similar base of knowledge" (Teigland 2003, p. 34). Teigland further notes that people working in the same occupation tend to develop similar values and vocabularies (Teigland 2003, p. 35), which facilitates knowledge transfer between them. Hence, they still fall under the definition of CoPs and Wenger et al. (2002) denote a whole chapter to their definition and management. Emphasizing the lack of face-to-face meetings Wenger et al. work out four areas that make it especially difficult to have a thriving distributed network of practice (or distributed Community of Practice as they name it). First, the geographic distances, but also possible time zone differences, decrease the feeling of the "presence" of the community for its members. Without actually seeing the others, it is for example difficult to understand if anyone at all or how many are benefiting from contributions. Along with the issue of distance comes the disadvantage of not having by-chance meetings e.g. at lunch or in the elevator, which decreases the possibility to maintain informal relations. Second, since distributed CoPs can attract large number of people, it is possible that they attract more than an individual can possibly know. This challenge has been discussed in section 3.4.2.3, by indicating that the feeling of "we" is an important property of a community such as a CoP. Wenger et al. still think that distributed Networks of Practice are a community even when a large number of individuals is involved. They acknowledge however that the ways of interaction change. Third, the affiliation to different parts of the company can be a hindrance to actively exchanging knowledge as the goals of the part of the company that an individual belongs to may be conflicting with those of another part of the company that other members of the distributed CoP are part of. Fourth, Wenger et al. note that a challenge arises when distributed CoPs cross cultural borders, as the proper norms of mutual engagement differ across cultures. To remedy this challenges somewhat, Wenger et. al. indicate that it is necessary to keep up "the rhythm" of the community by arranging regular face-to-face meetings.

Because of the missing ability to socialize in person and share knowledge as side-product of it, in a Distributed Network of Practice, interactions and consequently the knowledge that is shared is often rather explicit in comparison to the classical form of CoP that are (also) focusing on more tacit elements. This is why typically the members of Distributed Networks of Practice are also members of (local) CoPs and are acting as brokers. In addition, the members of Distributed Networks of Practice are typically only loosely connected. On the other hand, this kind of social connection is less intense, requires less effort from the individual and the degree of social restrictions on what is proper behavior is lower. Nevertheless, as was true for classical CoP, interactions between members are still rather dyadic and reciprocal in nature (Teigland 2003, p. 39).

Teigland (2003) notes that due to their characteristic of being distributed, intra-organizational Distributed Networks of Practice emerge differently than CoPs because physical closeness cannot be relied on. She indicates some possibilities that she found in her literature analysis that should facilitate the creation and maintenance of Distributed Networks of Practice like creating the opportunity for face-to-face meetings in inter-office transfers, organizational retreats or conferences, or multi-office projects.

Distributed Networks of Practice are also possible across the legal boundaries of a company. In academia, this is very common, taking special interest groups of associations like the ACM or GI as an example. In non-academic business, there is also similar trends to see, taking (special interest) groups in social networking services such as Xing as an example. While for academia there have been research endeavors to understand the interactions between researchers, for other settings, inter-organizational Networks of Practice have not found their way into published studies. It is hard to obtain the necessary data and get access to individuals (Teigland 2003, p. 47). However, it is safe to assume that inter-organizational Networks of Practice exhibit characteristics comparable to those of intra-organizational Networks of Practice. Differences lie in the extent of some of the characteristics. It is for example even more difficult to maintain a common understanding, shared values and norms when the individuals do not share the same company background. This likely leads to less identification with the network. Moreover, the willingness to share knowledge across legal boundaries of companies might also be hampered by company regulations concerning knowledge leaking to the outside, although this still happens and can even be beneficial for the company (Brown, Duguid 2001, p. 207).

3.4.5 Electronic Network of Practice

Electronic Networks of Practice consist of a geographically distributed group of individuals that are engaged in a shared practice. However, in contrast to Distributed Networks of Practice, the group of individuals can be very large, *virtually limitless in size*, the individuals are *loosely knit*, and may not know each other at all nor necessarily do they expect to ever meet face-to-face (Brown, Duguid 2001; Wasko, Faraj 2005). Electronic Networks of Practice enable the creation of *weak structural links* between thousands of individuals irrespective of the individuals' social status, racial demographics or geographic location, as those are “filtered” by the mediating technology in contrast to face-to-face interactions where those would possibly limit the number of interaction partners due to explicit or implicit prejudice. Electronic

Networks of Practice allow quick and effortless access to a broad source of expertise through a wide variety of knowledgeable individuals (Teigland 2003). Individuals in Electronic Networks of Practice share knowledge, solve each other's problems, and learn through posting questions and answering the ones of others and by debating issues relevant to the network's topic (Wasko, Faraj 2000, p. 161). Individuals benefit from these networks as they may tab into new information, expertise, and ideas that may not be available locally. Brown and Duguid (2001, p. 206) reason that those networks extend an individual's reach orthogonally across the company's value chain, and along common practice.

While the possible number of members of the network is limitless – in Electronic Networks of Practice, anyone interested in the group's topic may join – the individuals are relatively anonymous towards each other. Often not more than a freely chosen name, possible an e-mail address and some voluntarily disclosed information is visible to others in the network (Wasko, Faraj 2000; Teigland 2003). Referring to the mechanism of CoPs, this lack of information about the background of the others in the network severely limits the creation of a shared identity, common language and norms as well as trust, obligation, and social controls, which Wenger (1998) sees as pivotal for the community's stepwise improvements of the practice. Next to the relative anonymity, the media used in electronic networks of practice is *less rich*, i.e. text-based messages that are "sent" via bulletin boards, forums, e-mails or chat rooms are more impersonal than face-to-face meetings and offer less possibility to create shared meaning. Along the relative anonymity in networks and due to the fact that messages are typically open to anyone in the network, the ties that are built up between the individuals in the network are less of a dyadic nature but individuals form *weak ties with the entire network* (Teigland 2003).

As noted before, Thiedeke (2003) also draws the implication that social norms are much harder to enforce in purely virtual settings which also holds for social rules of direct reciprocity. As a result, if an individual receives help from another there is no direct pressure to later help this concrete individual in return. Instead Wasko and Faraj (2000, p. 169) note that " […] giving back to the community in return for help was by far the most cited reason for why people participate" in their study of an Electronic Network of Practice. Hence, the reciprocity is *generalized* from receiving likewise support from the same individual to receiving likewise support from (some arbitrary individual from) the network. Additionally, Wasko and Faraj (2000; 2005) found some more reasons in their study explaining why individuals contribute to Electronic Networks of Practice: it is challenging, helping to refine the own thinking; it brings enjoyment to solve problems; it feels good to be seen as expert; it increases one's own reputation, it helps in enhancing self-efficacy. As the resulting approach of this thesis will have similarities with Electronic Networks of Practice, all of these motivational aspects are possibly relevant.

3.4.6 Advantages

The different kinds of social constructs that were described so far have distinct advantages that render them suitable for KM initiatives in companies. In the following, the advantages of the three kinds of social constructs that were described are highlighted.

3.4.6.1 CoP

CoPs have become a quite popular measure especially in larger corporations in recent years. They promise distinct advantages. For example, Schütt (2003, p. 458) notes that one problem of companies is that relevant experts, that could be found using YPS or ERS, might leave their job position or the company. If an expert’s knowledge was not redundant in some way in the company, this leaves a gap in what the company needs. Schütt indicates that CoPs can be an effective measure against this kind of problem as CoPs foster a *shared set of expertise* and this way can more flexibly cope with an individual expert leaving the company. He further notes that CoPs can also be an intriguing source of information for the company’s management about which knowledge the individuals of the company posses and hence about the pool of knowledge of the company itself.

Next to the increased flexibility, one of the very reasons why CoPs emerge and are actively nurtured by companies is their ability to create new knowledge through incremental improvements in local work practices, best-practices that are spread or new methods that are established in response to new problems (Brown, Duguid 1991; Wenger et al. 2002). But not only fundamental changes to work practices are fostered. In CoPs individuals are acquainted to each other and have a good impression of who knows what. This allows them to ask for help in very specific situations, as they know whom to address. Further, being rooted in the *theory of situated learning* (Lave, Wenger 1991), CoPs also foster the learning experience of less seasoned employees from more seasoned employees in a natural way and by this develop professional skills within the company. Wenger and Snyder (2000) denote some more concrete advantages. For example, they indicate a case in which a CoP, first small and unknown, was the breeding ground for an entire new business segment of a consultancy. In another exemplary case, they state that the members of a CoP were successful in retaining another member by acquiring new projects for this member or by offering other incentives, like become a prestigious member of another CoP – a valuable HCM activity autonomously triggered by the CoP.

3.4.6.2 Distributed Networks of Practice

Distributed Networks of Practice have the advantage of extending the number of individuals that can be part of the community to a much larger number than is possible in “classical” CoPs. Naturally, some of the mechanisms of frequent face-to-face interactions typical for

CoPs do not hold in this setting. The lesser degree of informal interaction, which transfers much of the tacit knowledge of an individual as a by-product of casual conversation, decreases the amount of knowledge that can be spread and shifts the knowledge to more explicit forms that are more easily communicated in non-face-to-face interactions. Nevertheless, Distributed Networks of Practice, having less enforceable *social norms*, allow an individual to decide more freely which amount of time to dedicate to the CoP. This can extend to a state of only *lurking* in which the Distributed Networks of Practice member does not contribute himself but still the company as larger entity gains benefit from the individual's peripheral learning effects.

3.4.6.3 Electronic Networks of Practice

Electronic Networks of Practice extend the number of possible interaction partners to a virtual limitless number of individuals. Wasko and Faraj (2000) could identify distinct advantages from their survey among members of Electronic Networks of Practice. For example, posting questions to the community can be a fast way to get answers to complex problems. Considering the often large number of members in the Electronic Network of Practice it can be *much faster* than in (distributed) CoPs. Other members indicated that knowledge not available through other means is often found in Electronic Network of Practice. The possibility to learn from the mistakes of others and to *enhance one's own self-efficiency* but without the *social pressure* to do so is seen as another benefit.

3.4.7 Disadvantages

While offering clear advantages, there are also challenges of the three social constructs. In the following, those disadvantages are highlighted for each of the three kinds of social constructs that were described.

3.4.7.1 CoP

The active support and management of CoPs is problematic. Wenger et al. state that it is far from easy to build, sustain and properly integrate CoPs (Wenger, Snyder 2000, p. 140) and that especially the organic, informal nature makes them *resistant to supervision and interference*. To still try to nurture them is, in their words a "managerial paradox" (Wenger, Snyder 2000, p. 140). Other authors support this argument stressing that CoPs can hardly be created but the building can only be facilitated by e.g. providing communication and collaboration means (Petter et al. 2007, p. 7). While both give general recommendation, none of them gives concrete advice how to manage CoPs.

Another challenge for companies to implement CoPs lies in the "double assignment" of individuals, in formal work groups and CoPs. Once CoPs are obvious to the organization their legitimization becomes an issue. Work group managers may start to use their subordinates to push their agenda into the CoP for political reasons (North et al. 2000). In inter-organizational CoPs this can be more intense as the feeling of competition may outweigh the need and ability to collaboratively learn (Petter et al. 2007, p. 1).

North et al. as well as Wenger et al. state that CoPs will decide for themselves who is going to be a member and who is not. However, the first step to becoming a member is knowing that there is a suitable community, a challenge which they do not address. Similarly, when there is no suitable CoP but the company would benefit from one because many individuals share the passion to exchange their work-related knowledge, it is difficult to bring those individuals together – they do not (yet) know of each other and their potential to actively work together. Both, Wenger et al. and North et al do not adequately address this challenge. Further, both state that CoPs can grow to a significant size which sometimes renders splitting into sub-groups as a viable option. While certainly this evolves from within the CoP and its communication and interaction structure, still it is difficult to understand when and how to build sub-groups and especially for which topics given only the more general topic of the encompassing CoP.

Finally, there are problems intrinsic to the constituting characteristics of CoPs. They draw their major benefit from rather intimate *strong ties* between their members, which necessarily limits the number of members to a smaller number: The more members a CoP has, the less it is possible to stay in frequent contact with everyone and consequently the less intimate the relationship can be. Also, an essential characteristic of CoPs is their informal aspect, e.g. by-chance encounters and informal chats that convey much of the tacit knowledge. This is however only possible when physically meeting and hence CoPs are limited to groups of individuals that are co-located.

The problems of CoPs are summarized in Table 3-6.

Id	Description
Problem 1	CoPs are difficult to manage; they are a managerial paradox
Problem 2	May be misused for political reasons
Problem 3	Knowing about the existence of a community as an outsider is difficult
Problem 4	Creation of advantageous but not yet existing communities is difficult
Problem 5	No support for when and how to split into subgroups
Problem 6	CoPs are very limited in size
Problem 7	(classical) CoPs can only form where people physically meet

Table 3-6 ***Summary of problems that Communities of Practice face***
Source: Own Illustration

3.4.7.2 Distributed Networks of Practice

The limited ability to meet face-to-face and to have informal communication create a lesser *sense of community* than in “classical” CoPs. This is why a minimum number of face-to-face encounters per year is necessary and regular other events should be scheduled to keep up the “rhythm” of the community (Wenger et al. 2002). However, the confidence that can be built up in face-to-face encounters fades over time. When individuals disperse again to different locations within the company, the level of confidence among them declines, until they meet face-to-face again (Teigland 2003). Therefore, companies are faced with a substantial effort to arrange for regular meetings. The current emergence of *social software* in corporate settings can possibly reduce the effect of fading confidence as they can support the communication and the *feeling of presence* in a Distributed Networks of Practice (Petter et al. 2007, p. 3). Also, it is possible that social software, in addition or as a replacement of classical YPS, allows to better display one’s interest and therefore might facilitate finding individuals with common context (Richter, Koch 2008, p. 19). Nevertheless, Distributed Networks of Practice require dedicated resources, typically more than local CoPs, for travel and the effort of bridging time, geographic, and cultural barriers can be high. At the same time, Distributed Networks of Practice are less capable of transferring knowledge than local CoPs.

On another account, Distributed Networks of Practice have a harder time conveying their advantage. While in local CoPs individuals often have informal meetings and the barrier between casual chat and goal-oriented knowledge transfer becomes blurry, in distributed settings this is much harder. Especially in inter-organizational Networks of Practice this can be a serious challenge. It is therefore not surprising that Petter (2007, p. 8) found in his study that the most successful communities were the ones that were most successful in defining the advantage for each individual.

Individuals in Distributed Networks of Practice know each other and maintain *dyadic relationships*. While Wenger et al. (2002) showed that it is natural for any community to have members with varying involvement and activity, the personal acquaintances between members of any CoP puts some social pressure on members to be active. This decreases the possibility for lurking also in Distributed Networks of Practice. Although this might seem a good thing, Wenger et al. (2002) note that lurkers are important parts of CoPs as well as they learn by “active listening”. In Distributed Networks of Practice, this is less possible than in more anonymous forms of interaction such as in Electronic Networks of Practice.

The increased size and along with it the inability to keep abreast of all developments of its members creates another challenge for Distributed Networks of Practice. It becomes increasingly challenging to know which knowledge resides in it and additionally it is challenging to know who possess the relevant knowledge, i.e. it is hard to find the *right* person to contact (Qureshi et al. 2006). This is problematic as Distributed Networks of Practice – similar to

CoPs – a rely on *dyadic relationships* and *direct reciprocity* which is why it is normal to approach selected individuals – typically the ones that "owe oneself" or the ones that one would want to "owe a favor". On top of this, while CoPs in general and Distributed Networks of Practice in particular focus on progressing the knowledge of a community that is bound by the same profession or part of a profession, they also serve the purpose of offering mutual support among its members when advice is needed. However, the need for advice and more generally someone else's knowledge is rooted in the current work context of the one needing help (Qureshi et al. 2006). As a result, very detailed knowledge of who knows what is necessary – specific to the task and specific to the current context. The larger a community is and the less connected, as it is the case for Distributed Networks of Practice, the more challenging it is to know so detailed about the help someone is able to provide.

The initialization of Distributed Networks of Practice is a complex task. They are *emergent networks*, i.e. they cannot be mandated but evolve on their own while supporting and nurturing from the company is possible. In local CoPs this can happen by supporting informal meetings or creating spaces for by-change encounters. In Distributed Networks of Practice it is much more complex. How to identify the right persons that could initiate and maintain a lasting Distributed Networks of Practice is a complex question.

In addition, some of the problems of local CoPs remain for Distributed Networks of Practice. To join a Distributed Networks of Practice, even before being invited to the community, it is necessary to know about the existence of the Distributed Networks of Practice after all. Similarly, if there is no suitable Distributed Networks of Practice for a topic, but there would be individuals ready to set one up around a specific topic, the formation is not supported in any way. Finally, the number of members of a Distributed Networks of Practice, one of its distinct advantages, may create the problem of being less able to focus on topics, which is an indication for splitting the community into smaller sub-communities. However, there is no support for knowing when and how a split would be advantageous.

The problems of Distributed Networks of Practice are summarized in Table 3-7.

Id	Description
Problem 1	High effort for regular face-to-face meetings
Problem 2	Harder than in local CoPs to convey individual advantage with respect to effort
Problem 3	Less ability for lurkers to just learn than in less anonymous settings
Problem 4	Larger size creates problems of maintaining sufficient levels of know-who
Problem 5	Knowing about the existence of a community as an outsider is difficult
Problem 6	Creation of advantageous but not yet existing communities is difficult
Problem 7	No support for when and how to split into subgroups

Table 3-7 ***Summary of problems that Distributed Networks of Practice face***
Source: Own Illustration

3.4.7.3 Electronic Networks of Practice

Wasko and Faraj's (2000) survey results shed light on possible challenges of Electronic Networks of Practice. Firstly, individuals may feel too uncomfortable with their level of expertise to help others. This may also be the case when the individual would be objectively capable of helping but subjectively feels otherwise.

Further, Wasko and Faraj cite respondents of their survey that indicate that sometimes they are unwilling to respond to requests although they were able to, because *implicit or explicit norms* were not lived up to by the requester. An example was an individual who stated that he helps if there is a problem but he does not "do the others homework", i.e. help is only granted if the requester shows sufficient evidence of attempts to solve the problem alone. This includes using the right terminology that shows minimum level of acquaintance with the subject. So even when there is no community setting in its strict definition, some *shared language* and *basic norms* can also be found in Electronic Networks of Practice. While they also establish in mutual interactions of individuals, they are "hidden" in the previous interactions and there is no *legitimate peripheral participation* (Lave, Wenger 1991) as in local CoPs in which new members learn the rules during interactions with more established members.

Additionally, Wasko and Faraj (2000) could identify a problem in too large networks. In difference to Distributed Networks of Practice, it is not problematic that individuals do not know whom to address with requests, as questions are always posted without direct recipient and virtually to all members of the network at once. However, when Electronic Networks of Practice become very large, participants start to have difficulties to differentiate personally valuable message and the information contained in them from the many other messages that are less valuable for the individual. This applies for two different processes. Some members may want to find open requests to answer and this way help others in the network. It might however be difficult to find those requests that can be answered due to the large number of messages. Analogous to this challenge, others in the network may use it to educate themselves contributing no answers or to a lesser degree but browsing through previous interactions of others to learn. Since in Electronic Networks of Practice all messages are generally visible for all members, finding ones that are relevant for an individual's concrete situation and level of knowledge can be problematic. The sheer size along with the inability to know all or even some of the individuals in the network also creates challenges similar to the problems that arise in large companies. It becomes a challenge to know which knowledge is contained in the individuals of the network.

An illustrative example is depicted in Figure 3-13. It shows the number of messages in two forums of an Electronic Network of Practice over a period of 8 months. The left is concerned with SAP Dashboard Design (formally known as Xcelsius), a business intelligence visualization solution of SAP, the right one is concerned with data warehousing using SAP solutions, so both are related with respect to their content and it is likely that there are many users interested in both topics. On average, in the SAP Dashboard Design forum there are about 16 messages a day posted with peaks of up to 58 messages, in the data warehousing forum there are on average almost 28 messages a day with peaks of up to 79 messages. In combination an individual interested in both forums would be required to skim over roughly 44 messages on average with peaks of up to 121 messages per day. However, the Electronic Network of Practice – the SAP Community Network – has more than 230 forums. While certainly not all of them are relevant for everyone, it is obvious that staying up-to-date with "only" those messages that are possibly relevant for someone and scanning those can consume tremendous amounts of time.

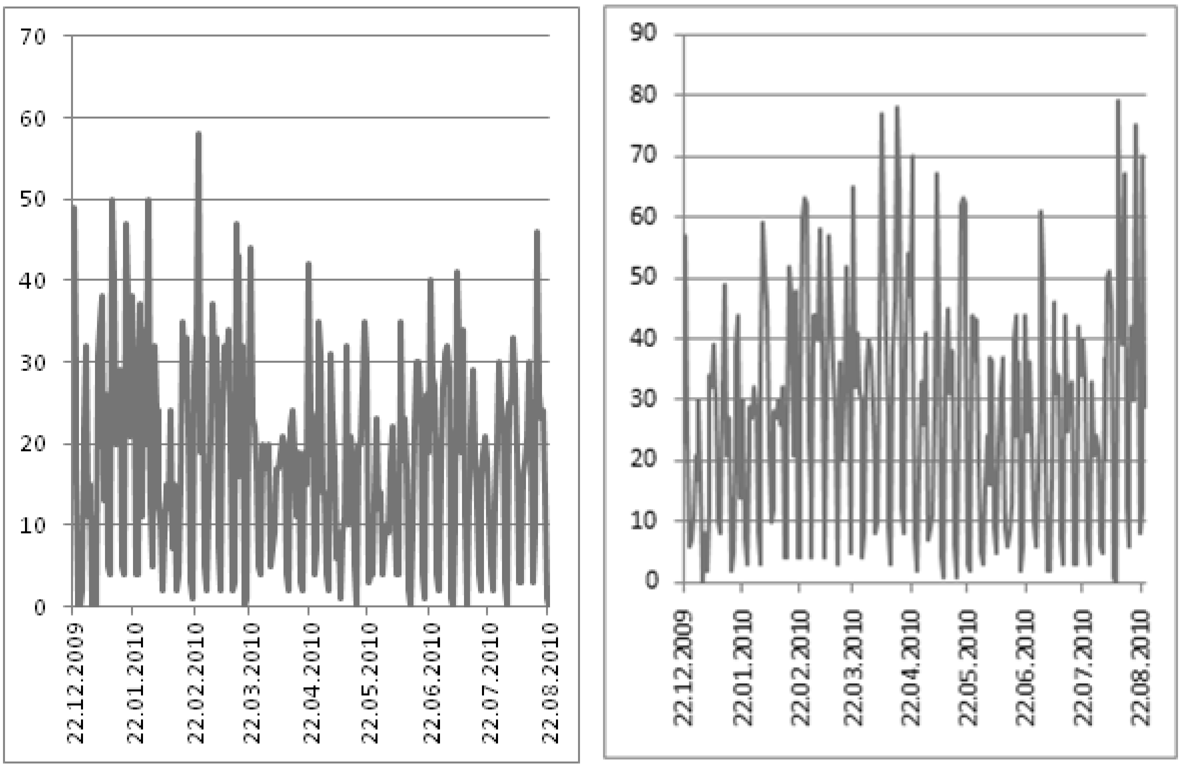


Figure 3-13 ***Number of messages in two forums of an Electronic Network of Practice focused on business intelligence solutions of SAP***
Source: Own illustration

In addition, some of the problems of CoPs and distributed Networks of Practice remain in Electronic Networks of Practice. It is still an effort for an individual to find Electronic Networks of Practice that could be beneficial for the individual. This is more complex than it seems, taking the SAP Community Network mentioned in the paragraph before as example. The more than 230 forums are identified by their name, so a rough-cut distinction of what might be relevant for an individual is possible, but to be sure the user would have to skim over (some of) the messages in the forums. Also still, there is no support for the creation of possibly advantageous networks that many could benefit from, but that lack the igniting spark. And, finally, the scaling of networks and their forums that they might feature can be problematic and the creation of subgroups might be necessary. Yet, there is no support of when and how to split into subgroups.

The problems of Electronic Networks of Practice are summarized in Table 3-8.

Id	Description
Problem 1	Limited ability of self-reflection may decrease contribution likelihood
Problem 2	Minimum level of shared language and norms apply but are not (openly) negotiated
Problem 3	Abundance of messages limits willingness to find requests than can be answered
Problem 4	Abundance of messages limits capability to learn from other's interactions
Problem 5	Knowing about the existence of a network as an outsider is difficult
Problem 6	Creation of advantageous but not yet existing networks is difficult
Problem 7	No support for when and how to split into subgroups

Table 3-8 ***Summary of problems that Electronic Networks of Practice face***
Source: Own Illustration

3.4.8 Summary

Similar to YPS and ERS, knowledge networks are set up to support the transfer of knowledge through mediating socialization processes. However, in contrast to these two approaches, knowledge networks focus on bringing together a group of individuals that interact with each other instead of supporting one-on-one interactions. This way, knowledge is transferred from one individual to one or more others. There is a *collective learning effect* and the resulting network of knowledge carriers is more robust against leaving experts. Additionally, new knowledge is created typically not by individuals but in discussions and interactions of groups of individuals. Therefore, the interaction allows increasing the collective knowledge – for each participant and the organization.

Different terms for similar concepts that deal with groups of individuals are used. This thesis chooses to use *knowledge networks* as most general term and analyzed (local) *Communities of Practice*, *Distributed Networks of Practice* and *Electronic Networks of Practice* as most relevant instances thereof.

Communities of practice (CoPs), in their original form, are groups of individuals that are determined by three interdependent structural components: *domain*, *community* and *practice*. The groundwork for CoPs can be found in the theory of *situated learning* which postulates that knowledge transfer is not a one-way "copying" from one individual to another but is actually a *social interaction* between the two that is situated in a certain context. CoPs are *emergent groups*, i.e. they form and evolve on their own. While a company can support and nurture CoPs, it cannot as such mandate them. They are different and additional to other forms of interaction such as work groups, projects, or business networks. To manage CoPs seven principles have been postulated (Wenger et al. 2002) and four dimensions of intervention have been identified (North et al. 2000). CoPs are communities of individuals and hence there is a "feeling of we". *Strong personal ties* among members are typical and hence a limited number of well-acquainted members, expectation of *direct reciprocity* and mutual (re-)definition of *shared norms and rules* are typical characteristics of CoPs.

With the emergence of computer-mediated communication the concept of CoPs was transferred into the "virtualized" world. An abundance of different terminology describes the phenomenon of communities that do not at all or only seldom meet face-to-face but still establish a form of community. However, the virtualized interaction changes some social aspects as well. The interactions are more anonymous than in face-to-face interactions, social norms are harder to enforce but the number of possible interaction partners can be extended to a large degree. Two forms of groups that rely on virtualized interactions were analyzed: Distributed Networks of Practice and Electronic Networks of Practice.

Distributed Networks of Practice contain individuals with similar profession or tasks that are dispersed in different geographical locations. There are *intra-* and *inter-organizational* Distributed Networks of Practice depending on whether the members come from different organizations or are within the same legal boundaries. In Distributed Networks of Practice, individuals often maintain weaker ties among each other than in local CoPs. However, *direct reciprocity* and *dyadic relationships* are still core aspects of distributed CoPs and the personal acquaintance between the members is kept as much alive as possible by arranging regular face-to-face meetings. The largely increased number of possible interaction partners however may increase the challenge to keep an overview of all participants.

Electronic Networks of Practice are groups of individuals that are geographically distributed, engaged in a shared practice, but, in contrast to Distributed Networks of Practice, can be very large, *virtually limitless* in size, the individuals are *loosely knit*, and members may not know each other at all nor necessarily do they expect to ever meet face-to-face. The ties that are built up between the individuals in the network are less of a dyadic nature but individuals form *weak ties* with the entire network. Hence, there is also no expectation of direct reciprocity but instead individuals expect *generalized reciprocity*, i.e. they expect receiving likewise support not from the same individual but receiving likewise support from (some arbitrary individual from) the network.

Analyzing the areas where the different forms of knowledge networks face challenges one can conclude some learning for creating this thesis' KM approach. First, in all three forms of knowledge networks individuals are faced with the initial challenge of knowing about the existence of a suitable knowledge network. In local CoPs this implies knowing the right persons within CoPs that can tell about its existence. In virtualized settings such as in Distributed Networks of Practice and Electronic Networks of Practice this includes browsing through directories of possible knowledge networks and determining if they are suitable or not. It would be advantageous to decrease this effort for potential participants. Second, similarly, in all knowledge networks the formation of new groups of individuals to exchange knowledge on a certain topic that is of interest to all potential participants, is not supported. In companies, the formation might be governed by top-down installation of knowledge networks and then supporting them. However, there is no support for the emergent, bottom-up formation of knowledge networks, which is often the way in which social interactions start to form longer lasting installations (Richter, Koch 2008, p. 20). Third, while reaching many possible interaction partners offers the advantage of a higher chance that someone can and will help and an increased number of different perspectives on a challenge are possible, too large numbers of individuals can induce a lack of overview, i.e. *information overload*. In all forms of knowledge networks, this problem is addressed by splitting into smaller, more manageable and focused subgroups of individuals. However, besides personal relationships among members that would suggest obvious possibilities how to split a knowledge network, there should be support or suggestions of how to split, especially in larger knowledge networks where interpersonal relationships are weaker. Fourth, while learning through the knowledge network is one of the goals anyway, in Distributed Networks of Practice and Electronic Networks of Practice the number of interactions that can lead to *information overload* can hamper the learning. There should be means for supporting the learning of each individual without overly burdensome mechanisms in place. Fifth, especially in Distributed Networks of Practice, there is a challenge of knowing who knows what in order to address the right persons. Nevertheless, also in Electronic Networks of Practice there is, a somewhat lower, *mental effort* to know where to post requests to. There should be support for the individual to limit this mental effort as much as possible. Sixth, there should be *non-invasive means* to foster the contribution of individuals. While in local CoPs *social pressure* and *expectations of reciprocity* ensure mutual contribution, this applies less so in Distributed Networks of Practice and virtually not at all in Electronic Networks of Practice. However, individuals are more likely to contribute, when they *feel competent* to do so and when (implicit or explicit) *norms* of the requestor are respected. So, both aspects should be fostered to support contribution behavior. Seventh, and in relation to the statement before, it should be obvious for each individual why to participate. There can be different motives, like *gaining reputation* and hence status, *intellectual challenge* and *joy of solving problems*, or *expectation of (direct or generalized) reciprocity*, that should be fostered.

The conclusions drawn are summarized in Table 3-9.

Id	Description
Conclusion 1	Awareness of individuals of existing possibilities for knowledge exchange should be supported
Conclusion 2	Creation of new possibilities for knowledge exchange should be supported
Conclusion 3	Support for creation of more focused fragments of existing possibilities for knowledge exchange should be supported
Conclusion 4	Learning of larger group should be fostered
Conclusion 5	Individual should be relieved as much as possible from the task of maintaining know-who
Conclusion 6	The likelihood of an individual contributing should be increased
Conclusion 7	Obvious benefit for individual participation should be made clear

Table 3-9 ***Summary of conclusions from analyzing expert networks***
Source: Own Illustration

3.5 Summary and Discussion

In this chapter, the current state of the art of approaches that support knowledge exchange by building upon socialization processes was examined. The choice of inspected approaches was derived. For this purpose, the kinds of KM processes that this thesis wants to support are defined and by relying on existing systematizations of approaches YPS, ERS, and knowledge networks are selected as relevant approaches that are most similar to this thesis' approach. Each is analyzed in turn to shed light on their principle ideas, advantages, and disadvantages. One can conclude that, while of course tailored to different needs, each approach faces some distinct challenges that were described and out of which conclusions were drawn for a possible KM approach that tries to address these downsides. Those conclusions are described in the following and are summarized in Table 3-10.

The less than expected adoption of YPS showed that any *additional effort* for KM approaches should be as low as possible for the individual (Conclusion 1). Further, it showed that KM approaches that rely on *models of competence* such as YPS as well as some ERS should support the company-specific definition of what competencies are relevant and how to define a suitable model (Conclusion 2). Moreover, the setup of a relation of an *individual's model of competences* with respect to the global competence model should be supported (Conclusion 3). In knowledge networks, more specifically in CoPs and Distributed Networks of Practice the previous interactions between individuals replace the need to build up a competence model. By interacting the individuals learn about who knows what. However, if Distributed Networks of Practice grow too large the problem might arise again. In Electronic Networks of Practice, the third form of knowledge networks that was addressed, there is no direct need to know who knows what as requests are posted virtually to all members simultaneously. Another way to increase the benefit of KM approaches is by relieving the individual from switching the (mental) context, i.e. the KM approach should ideally be *embedded* directly in the individual's *current work context* (Conclusion 4). In knowledge networks, especially in CoPs and also to a lesser degree in Distributed Networks of Practice this conclusion only partly applies. The spectrum of valuable topics for discussion is broader there and individuals discuss more general characteristics of their work. In Electronic Networks of Practice and in everyday task situations, in which YPS and ERS are often utilized, the statement however applies, as more specific knowledge chunks become relevant and the time between request and response tends to be smaller.

YPS and ERS exhibit a downside in that they lack support for *more generalized* learning beside one-on-one interactions (Conclusion 5), i.e. there is only knowledge transfer and mutual learning between two individuals. In knowledge networks the learning is spread across the community by means of regular interactions in the case of CoPs and Distributed Networks of Practice or by posting to the general public in Electronic Networks of Practice.

Further, the analysis showed that the pursued goal to indicate the best individual with respect to a current need does not always have to be the most expert individual in the respective field. The possibility to actually interact with a more experienced individual and along this line the ability to communicate on a common basis is relevant as well (Conclusion 6). Both factors determine the *right* interaction partner. In CoPs and Distributed Networks of Practice the common basis can be found in *shared norms* and *rules* that have been mutually and continuously negotiated. In Electronic Networks of Practice this generally does not apply as no individual is addressed personally. However, also in Electronic Networks of Practice some common grounds with respect to rules and norms are established. In settings where interaction partners are typically less acquainted, like when seeking advice for a current problem using YPS or ERS, the *overlap in context*, *previous history*, and *shared language* is relevant.

When utilizing knowledge networks, especially Distributed Networks of Practice and Electronic Networks of Practice, it may be challenging for individuals to find already existing knowledge networks and hence the individual should be proactively supported to *find suitable ones* (Conclusion 7). Neither YPS nor ERS aim at bringing together groups of people, rather they aim at supporting the establishment of dyadic relationships. This conclusion hence does not apply to them. Nevertheless, the facilities of recommending experts might be helpful for recommending suitable communities, i.e. ERS might support knowledge networks.

Additionally, when there are no suitable knowledge networks, the *formation of new ones* may be beneficial if there are sufficient numbers of individuals interested in a topic. However, the build-up of those emergent groups should be facilitated – individuals might not know that others could be interested (Conclusion 8). Again YPS and ERS do not aim at creating networks of individuals and hence the conclusion does not apply to them. Nevertheless, as stated before, the formation of like-minded individuals into a network could be facilitated by ERS.

A similar challenge lies in the splitting of “too large” knowledge networks. Within a community or network the number of participants can grow to an extent that makes it increasingly hard to keep an overview and to keep the discussions focused enough to create benefit for all. In this case, the formation of sub-groups within the community or network is beneficial. However, the *split into sub-groups* should be supported somehow to trigger the process and to facilitate the setup of new sub-identities (Conclusion 9). As with the two conclusions before, YPS and ERS pursue a different goal and hence this does not apply for them. Again, the facilities offered by ERS might be helpful in sub-group formation.

Finally, the investigation of the three approaches for fostering knowledge exchange showed that individuals should be *encouraged to engage* in active knowledge exchange. In CoPs where direct social interaction is the fuel for knowledge exchange, the enforceability of social norms and rules ensures suitable knowledge exchange. In YPS that mainly focus on reflecting who is expert in which domain, the mechanisms for encouraging knowledge exchange lie within the characteristics of the individual searching for advice. However, in ERS as well as in Distributed Networks of Practice to some degree and in Electronic Networks of Practice to a larger degree, the computer-mediated interactions allow to and should be used to encourage contribution by adequately utilizing the different features of (partly) *anonymous*, *asynchronous* communication (Conclusion 10).

Knowing now where there is room for improvement, the next chapter analyzes which general aspects have to kept in mind when designing this thesis’ KM approach.

Id	Approach	Description
Conclusion 1	YPS / ERS / Knowledge networks	Effort for the individual must be reduced as much as possible, with respect to expected benefit
Conclusion 2	YPS / ERS	Means for supporting the maintenance of (implicit or explicit) competency models is necessary
Conclusion 3	YPS / ERS / knowledge networks (only Distributed Networks of Practice)	Proof for competency or skill should amend self-reflection, Expertness should be determined building upon more direct evidence such as users’ typical tasks
Conclusion 4	YPS / ERS / knowledge networks (not CoPs and Distributed Networks of Practice only to a limited degree)	Means for connecting knowledge seeker and knowledge bearer should support the user in a contextualized way, taking into account social aspects

Id	**Approach**	**Description**
Conclusion 5	YPS / ERS / Knowledge networks (already possible, but could be improved)	Means to increase the benefit from the interaction to more than just the two partners are desirable
Conclusion 6	YPS / ERS	Absolute level of expertness should not be used alone to recommend “experts”; shared mental models should be considered also
Conclusion 7	Knowledge networks	An individual’s awareness of existing possibilities for knowledge exchange should be supported
Conclusion 8	Knowledge networks	Creation of new possibilities for knowledge exchange should be supported
Conclusion 9	Knowledge networks	Support for creation of more focused fragments of existing possibilities for knowledge exchange should be supported
Conclusion 10	Knowledge networks (CoPs only very limited) / ERS	The likelihood of an individual contributing should be fostered

Table 3-10 ***Summary of conclusions from the state of the art research of expert mediation approaches***
Source: Own Illustration

4 Success Factors for Knowledge Management Projects

"Langfristig ist man nur erfolgreich, wenn man weiß, warum man erfolgreich ist!"
Freely translated "In the long run, you will only be successful, if you know why you are successful!"
Rupert Lay (born 1929), German philosopher

This chapter concerns itself with factors and requirements that need to be taken into account and addressed adequately to ensure that this thesis' KM approach, will be as beneficial as possible. To this end, first, the concept of critical success factors is introduced and a glimpse at what the definition of success of KM projects could be is given. Next, the search for influential factors that determine the success of KM projects in general is described. Relevant critical success factors from the available literature are extracted following a well-accepted literature review methodology. This abstract view on important aspects is then extended in the next subchapter. As this thesis' KM approach will rely on IT support, the subsequent sub-chapter investigates which aspects have to be taken into account to make IT systems that support KM successful. Again, a literature review will be the method of choice for extracting the relevant factors, this time focusing on published KM case studies that concern themselves with those kinds of systems that were analyzed in chapter three. The chapter ends with conclusions that are, next to the result of chapter three, used as a basis for the development of this thesis' KM approach in chapter five.

4.1 Preliminary Considerations

Before describing the methodology for extracting relevant factors from the current literature[3] and the results that could be obtained, first the basis of the analysis is described shortly. In the next two sections, the concept of critical success factors are explained and how they influence success. In order to use them, it is necessary to understand what success is, or more precisely when the object of interest – the KM project – is seen as success. As will be seen later, this is a still unresolved issue as the results of KM projects are not directly measurable. However, in the analyzed literature, some indications of success were used of which the most common one is adopted for this thesis, while explicitly stating that it is open research to determine a more suitable definition of success of KM projects.

[3] The following results build upon results obtained during the diploma thesis Slavtchev (2009) that was designed, initiated and guided by the author

4.1.1 Critical Success Factors

Owing to the very complex interplay of many contextual factors and many possible influences that are present in a company, the only viable option to learn about what influenced success is by relying on retrospective analysis. If a larger number of case studies can be inspected, it is possible to extract those factors that contribute the most to the success or failure of an initiative. The concept of focusing on the most important factors that influence an initiative's success is also called determination of *Critical Success Factors* (CSFs), an approach first introduced by Daniel (1961). This approach became more popular throughout different areas of management research after its refinement by Rockart (1979).

Briefly, CSFs are a limited number of key aspects that have to be fulfilled to a satisfactory level to ensure success. Those are the aspects where "things must go right" for the overall goal to be archived (Rockart 1979, p. 85; Bullen, Rockart 1986, p. 285). Knowing the CSFs for a certain area such as KM initiatives allows the responsible persons to focus their attention, which increases the effectiveness of the initiative. Because the design of a new KM concept is the goal of this thesis, it is necessary to collect, aggregate and analyze the relevant CSFs for KM initiatives in order to later on, design the concept along them.

4.1.2 Definition of Success in KM Initiatives

CSFs try to determine what influences the success of a KM initiative. However, before the influence can be determined it is necessary to understand what *successful* actually means, i.e. the question arises: When exactly can a KM project be specified or labeled as successful?

This is a non-trivial question, because most of the effects of KM initiatives are not or only indirectly quantifiable. Is the contract for a new large project a result of a successful KM initiative that allowed to leverage the company's knowledge assets better or is it rather attributable to a marketing campaign that was launched a while ago or to the redesign of the company's product. The direct influence is not obvious. Furthermore, if the KM initiative did contribute to a company's success after all, how large was its contribution? Again, this is often not quantifiable.

While there has been much research on the area of measuring KM performance there is no measure or set of measures that enjoy full support from everyone. Therefore, this thesis adopts a list of success measures as proposed by Davenport et al. (1998). In the analyzed literature, it has been cited frequently which suggests large support considering the current state of the KM community' knowledge base.

Seeing KM initiatives as specific case of general business change projects, Davenport suggests four success indicators that he derived from indicators for other areas. KM initiatives are seen as successful if all or at least some of the following can be observed (Davenport et al. 1998, p. 48):

- *Growth in the resources* granted to the project. This includes employees, money, facilities, etc. and is based on the assumption that resources are only provided for initiatives that have a positive impact and contribute value to the organization.
- *Growth in the volume of knowledge content* within the system and increased usage. Exemplary measures can be the number of (useful) documents, access rates to provided content, and participation in the case of discussion-oriented communities.
- Likelihood that the project would *survive without the support of particular individuals* that are often the drivers behind the initiative. Essentially, the main question here is, if it is an isolated individualistic project or a broader initiative accepted and lived by a large part of the organization.
- Evidence of *financial return* or a clear cut *business case*. Financial impact does not need to be rigorously specified since it often cannot be solely attributed to the KM program itself and thus calculated precisely. Examples for direct measurable returns range from KM initiatives having their own profit center, initiatives that focus on patent licensing and create quantifiable results, or in more simple cases work time savings and improved operations calculations.

4.2 General Success Factors for KM initiatives

In its history, KM has seen a considerable number of failures. Initiatives were introduced with large effort but did not deliver the benefit that was expected (see for example Braganza, Möllenkramer 2002; Chua, Lam 2005). However, other initiatives led to well-received improvements in the company's business (see Bellmann et al. 2002 for a compilation of cases). Naturally, the question arises which factors determine whether an initiative would be successful or not. In the following the approach for determining those factors – the CSFs – is described, before an overview of the results is given followed by a detailed explanation of each identified CSF.

4.2.1 Approach for Determining CSF for KM Initiatives

In the KM context, there have been a number of approaches to determined CSFs. To integrate their results a *literature review* on those factors was conducted (Schmidl et al. 2011a). The literature review followed the methodology proposed by Baker (2000), Webster and Watson (2002), Torraco (2005) and Fettke (2006). The search for relevant publications started with well-established literature on foundations of KM including among others (Davenport, Prusak 2000b; Nonaka 2001; North 2005; Probst et al. 2006; Lehner 2008) and also by searching for KM review articles of respected conferences and journals. The goal of this initial step was to find the right terms to use for the tailored search for KM CSFs.

The literature review uses the iterative *concept-centric approach* described by Webster and Watson (2002). In comparison to other literature review methodologies, the concept-centric approach categorizes and structures the literature around identified phenomena and concepts and does not focus on compiling mere summaries, which is often the case in e.g. author-centric reviews (Webster, Watson 2002). The main resources for identifying the available literature included:

- Electronic libraries such as the Gateway Bayern and the local university's Online Public Access Catalog (OPAC) especially for monographs.
- The DocumentWEB and DBIS electronic journal and magazine system to access journal articles and magazines. The main keywords applied for searching were "Knowledge Management" and "Wissensmanagement".
- The EBSCO Host database was used to identify journal articles and additionally utilized to perform forward and backward searches. The search terms consisted of Critical/Key/Success Factors + Knowledge Management/KM, Erfolgsfaktoren + Wissensmanagement/WM, Sins/Barriers + Knowledge Management/KM, Barrieren + Wissensmanagement/WM, and literature review. They were applied in all possible combinations.
- The search engine Google and its literature search engine GoogleScholar were also used with the same keywords to identify additional journal and conferences papers.

After analyzing the titles and abstracts and performing *forward and backward search*, more than 180 sources were downloaded or physically acquired. Once an article was identified in an accessible journal, a systematic *spanning search* was conducted in order to identify other suitable sources within the rest of the volumes and issues of the same journal. Some of the more widely known journals and conference proceedings included:

- Journals: MIS Quarterly, Journal of Computer Information Systems, California Management Review, Harvard Business Review, IBM Systems Journal, Journal of Knowledge Management, Knowledge & Process Management.
- Conferences: Professionelles Wissensmanagement – Erfahrungen und Visionen (2003-2009), International Scientific Conference Knowledge-Based Economy (2005), First World Summit on the Knowledge Society (2008).

After reading the identified set of publications in more depth, the ones that did not contribute to this thesis goal of identifying CSFs for KM initiative were filtered out, which left 43 relevant articles.

4.2.2 Overview of the literature analysis result

The goal of the literature review was to identify the CSF that influence KM initiatives' success. While all analyzed publications discuss success factors, the nine publications shown in Table 4-1 are most often cited in KM-related publications and also cover the remaining publications' proposed success factors. Therefore, one can regard them as central and the further analysis focuses on these publications.

Citation	Title
Skyrme and Amidon (1997)	The Knowledge Agenda
Davenport et al. (1998)	Successful Knowledge Management Projects
Liebowitz (1999)	Key Ingredients to the Success of an Organization's Knowledge Management Strategy
American Productivity & Quality Center (2000)	Executive Summary - Successfully Implementing Knowledge Management
Holsapple and Joshi (2000)	An Investigation of Factors that influence the Management of Knowledge in Organizations
Hasanali (2002)	Critical Success Factors of Knowledge Management
Chourides et al. (2003)	Excellence in KM: An Empirical Study to Identify Critical Factors and Performance Measures
Wong (2005)	Critical Success Factors for Implementing Knowledge Management in Small and Medium Enterprises
Hung et al. (2005)	Critical Factors in Adopting a Knowledge Management System for the Pharmaceutical Industry

Table 4-1 ***Main sources for the review of CSF articles***
Source: Own Illustration

The selected publications vary with respect to their research objective, applied methodology, and in their results. They range from pure qualitative and pure quantitative studies, to studies that use a combined approach by succeeding an initial qualitative research phase by an empirical interview or questionnaire-based evaluation. The studies were conducted either by professionals who have prior working experience in KM or by researchers that focus on CSF analysis for KM initiatives. Table 4-2 displays an overview of the publications' research objectives, applied research methodology, and the results that were obtained in the publication.

Author(s)	Research Objective	Research Methodology	Results
Skyrme and Amidon (1997)	Presentation of state-of-practice in KM and identification of CSF	Case study research of leading companies in KM	Seven main factors, other factors are only shortly addressed
Davenport et al. (1998)	Identification of the characteristical factors that make KM projects successful	Exploratory case study with focus on projects as research unit. A total of 31 projects in 24 companies were examined	Eight success factors
Liebowitz (1999)	Identification of the "ingredients" for successful KM	Review of lessons learned from early adopters of KM	Six core ingredients, various other are shortly addressed
American Productivity & Quality Center (2000)	Provide a Best-Practices report for successful KM implementation	Case study research supported by a consortium of 49 companies. Ten were chosen to participate in the study	Six success factors
Holsapple and Joshi (2000)	Development of a descriptive framework for success factors of KM	Literature review and a Delphi Study including KM researchers and practitioners	Three main success-influencing categories. Four CSFs within the managerial dimension
Hasanali (2002)	Presentation of CSFs for KM	Practical experience in the consulting industry	Five main categories with several sub-items
Chourides et al. (2003)	Derivation of CSFs and performance measures for KM initiatives	Literature review and longitudinal interviews with key staff in eight companies	A range of factors divided in five functional areas
Wong (2005)	Provide an integrative perspective of CSFs for implementing KM in the SME sector	Literature review and empirical assessment based on questionnaires. Participants comprised academics, practitioners and consultants	Eleven CSFs
Hung et al. (2005)	Discussion of CSFs involved in implementing a KMS	Literature review and empirical questionnaire-based research with 98 participants	Ten CSFs were identified

Table 4-2 ***Research Objective, Research Method and Result of the nine most relevant CSF publications***
Source: Own Illustration

Extracting the nine publications' proposed success factors, it is possible to distinguish a total of twelve success factors. Most of the authors use different, but similar terms and concepts when defining and referring to their CSFs. Whenever different terminology was used for the same concept this thesis' analysis adopts the most frequently used terminology. The CSFs are:

- Strategy
- Management Leadership and Support
- Organizational Culture
- Infrastructure and Roles
- Measurement
- Processes
- Knowledge Structure and Ontologies
- Information Technology and Knowledge Management Systems
- Motivation
- Training
- Human Capital Management
- Resources

Table 4-3 and Table 4-4 show the mapping of the terminology adopted in this thesis to the concepts used in the nine publications included in the analysis.

In the following, the detailed interpretation of the twelve identified CSFs is addressed. Each interpretation is summarized from the explications of the nine primary sources but amended with further detail by looking at related publications that were determined in the first phase of the literature review.

Author(s)	Strategy	Management Leadership & Support	Organizational Culture	Infrastructure & Roles	Measurement	Resources
Skyrme and Amidon (1997)	A compelling vision and architecture; Strong link to business imperative	Knowledge leadership (Top Management)	A knowledge-creating and sharing culture	Identifiable promoters and leaders		
Davenport et al. (1998)		Clear language and purpose; Senior management support	Knowledge friendly culture	Organizational infrastructure	Link to economic performance or industry value	
Liebowitz (1999)	KM strategy	Senior leadership support	A supportive culture	CKO or equivalent and KM infrastructure		
American Productivity & Quality Center (2000)	Strategy	Leadership	Culture	Organizational roles	Measuring and Evolving	
Holsapple and Joshi (2000)		Leadership			Measurement	Resources
Hasanali (2002)		Leadership	Culture	Structure, roles and responsibilities	Measurement	
Chourides et al. (2003)	Strategy	Leadership commitment	Culture	Organizational conditions	Measurement	
Wong (2005)	Strategy and purpose	Management Leadership & Support	Culture	Organizational infrastructure	Measurement	Resources
Hung et al. (2005)		Senior management leadership and commitment	Trusting and open organizational culture; employee involvement and empowerment; trustworthy teamwork		Performance measurement; Benchmarking	

Table 4-3 Mapping of the concepts used in the nine most important CSF articles to the overall identified twelve CSFs - Part 1
Source: Own Illustration

Author(s)	Processes	Knowledge Structure & Ontologies	Information Technology & KMSs	Motivation	Training	HCM
Skyrme and Amidon (1997)	Systematic organizational knowledge processes		A well develop technology infrastructure		Continuous learning	
Davenport et al. (1998)	Multiple channels for knowledge transfer	Standard and flexible knowledge structure; Clear language	Technical infrastructure	Change in motivational practices		
Liebowitz (1999)		Knowledge ontologies and repositories	KM systems and tools	Incentives to encourage knowledge sharing		
American Productivity & Quality Center (2000)			Technology		KM Training	
Holsapple and Joshi (2000)	Control Coordination					
Hasanali (2002)			IT infrastructure			
Chourides et al. (2003)			IT	Motivation and rewards	Training	HCM
Wong (2005)	Processes and activities		IT	Motivational aids	Training and education	HCM
Hung et al. (2005)		Knowledge structure	IS infrastructure		Employee Training	

Table 4-4 ***Mapping of the concepts used in the nine most important CSF articles to the overall identified twelve CSFs - Part 2***
Source: Own Illustration

4.2.3 Critical Success Factor 1 – Strategy

A clearly communicated and well-planned KM strategy is one of the central means for achieving successful KM initiatives (Liebowitz 1999; Du Plessis 2007). There is common agreement that the KM strategy needs to be linked closely to a company's business strategy.

In a survey of KM practices, Hansen et al. (1999) found that companies succeeding in their KM endeavors specifically *align* their KM strategy to their competitive strategy, and the type of products and services they offer. The authors argue that competitive strategy must drive KM strategy, and suggest as first step that firms should determine the basis of their value creation by answering the following three questions:

- Do we offer standardized or customized products?
- Do we have a mature or innovative product?
- Do we mainly rely on explicit or tacit knowledge to solve problems?

Very generally, KM strategies can be classified in one of two broad categories: The *codification strategy* and the *personalization strategy* that are schematically also depicted in .

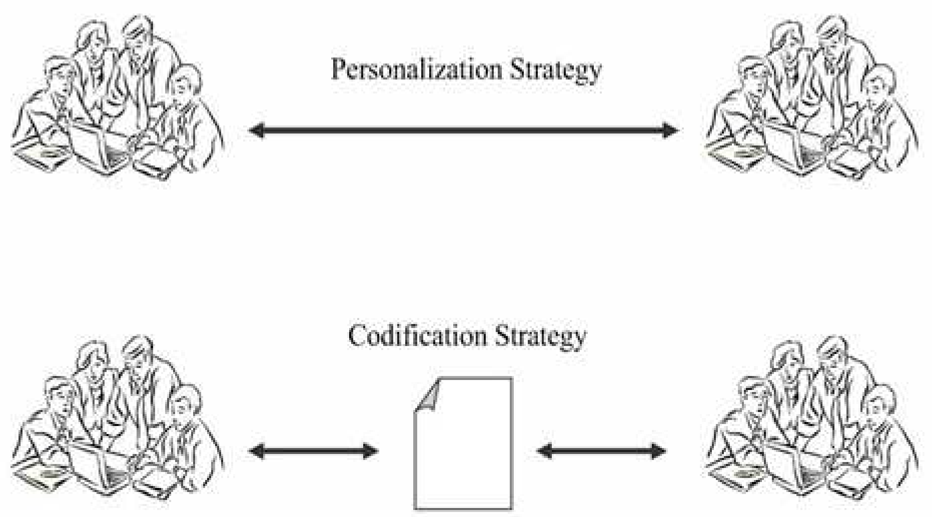


Figure 4-1 The two general KM strategies
Source: adapted from (Hansen et al. 1999; Krcmar 2010, p. 633)

The *codification strategy*, also often called *people-to-documents* strategy, predominantly focuses on *explicit* knowledge, and uses IT and Knowledge Management Systems (KMS) as tools to store, i.e. codify and share it. This strategy is advantageous if quality assurance plays a larger role than flexibility and if the quick and widespread dissemination of explicit knowledge promises large benefits for a company. Further, this strategy allows independence from knowledgeable individual as their (explicit) knowledge is captured and can be reused without relying on this particular individual. The disadvantage of this strategy lies in its inability to transport tacit knowledge. Additionally, knowledge, even if stored in an KMS, almost always is stored in a context-dependant way, i.e. the consumer of the knowledge base needs some implicit knowledge such as the corporate jargon, to understand the contents of the data store. If not explicit knowledge, but problem solving support is sought, this strategy can only give limited support. For example, the ability to address a new challenge that has not been encountered yet, might be supported by a knowledgeable co-worker but not (directly) by a document that has been created by this co-worker. Finally, the creation of knowledge bases without previous demand, which is typical for this strategy, might result in the spending of resources without knowing if they will ever be used (Hansen et al. 1999; Krcmar 2010).

The *personalization strategy*, also often called *person-to-person strategy* capitalizes on expert economics. It relies on and tries to foster direct communication and networks of people sharing *situational tacit* knowledge and their personal experiences. IT and KMSs are mostly used to categorize employee skills and connect knowledgeable personnel to knowledge seekers. The advantages of this strategy lie in the *problem-oriented* and/or *task-oriented* transfer of knowledge and provision of expert advice (Hansen et al. 1999; Krcmar 2010). The disadvantage of the person-to-person strategy lies in its restricted scalability and the inherent dependence on individuals and their ability to effectively transfer knowledge (Hansen et al. 1999; Krcmar 2010). In addition, knowledge transfer in this strategy often is a one-on-one activity, making it complex to spread once acquired knowledge quickly throughout the company.

The two strategies are stereotypes and should not be considered as mutually exclusive. Also focusing on one while entirely neglecting the other is not advisable, but rather a combination of both is advantageous. Hansen et al. (1999, pp. 112–114) confirm this observation and found that companies who use knowledge effectively pursue one strategy predominantly and use the second to support the first. They think that striving for perfecting in both simultaneously is not the way to go and rather suggest an 80-20 split. This rule is however quite simplified as Koenig (2004) states. He contradicts the 80-20 splitting rule by analyzing the pharmaceutical industry, where he found that the more successful pharmaceutical companies deliberately adopted an even split of resources on both strategies. Based on these findings he argues that the 80-20 split, or 20-80 split respectively, resemble the practical limits of the range in which an appropriate KM strategy mix is to be found. Moreover, the correct balance is likely to differ within different functions, the individual units of the organization, and even among single processes of the same department. As such, the decision of which KM strategy to adopt and to which extend needs to be examined on a much finer level than what Hansen et al. (1999) proposed.

Another argument concerning an advantageous link between KM strategy and business strategy is brought up by Zack (1999). He argues that not only does the business strategy determine which KM strategy to choose, but also it can be the case that the KM strategy can adapt the current business strategy. If the KM and the business strategy are not fully in line, gaps are the result. Figure 4-2 illustrates the two general gaps that can exist in a company. There can be a gap between what firms *can do* and what they *must do* according to their business strategy, i.e. a strategic gap. In addition, there can be a gap between what firms *do know* and what they *must know*, i.e. a knowledge gap. In this second case, if the company knows more than it must, it can exploit the "extra" knowledge to adapt the business strategy and potentially create a competitive advantage.

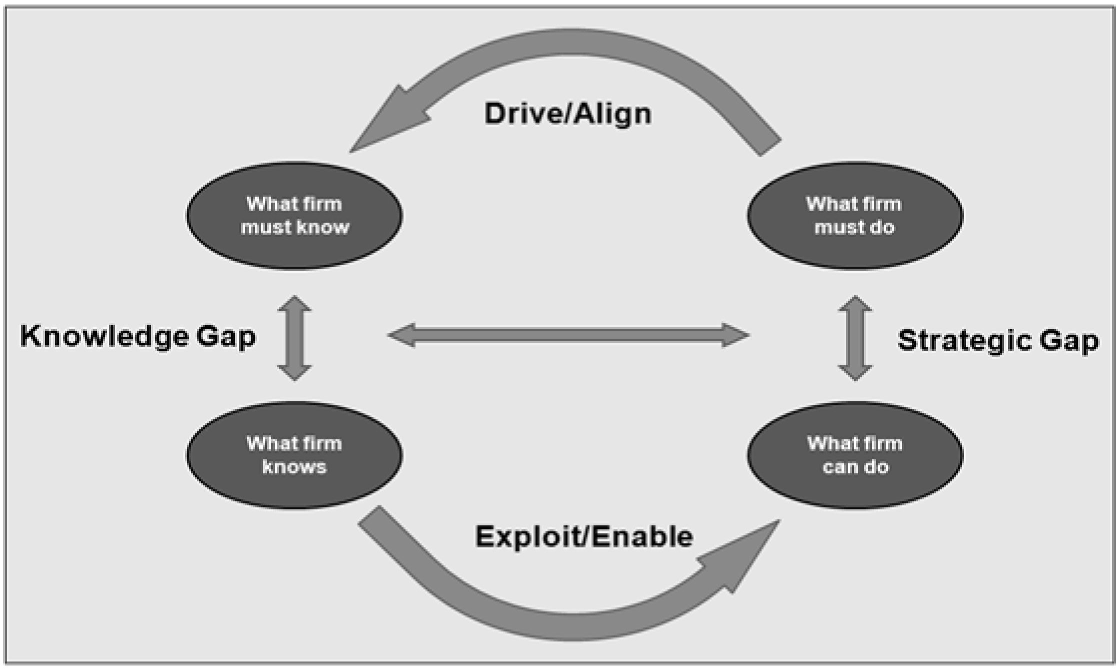


Figure 4-2 ***Relation between KM Strategy and Business Strategy***
Source: Adapted from (Zack 1999, p. 136)

Summarizing the discussion, there are two general approaches to KM, the first one relies on codifying explicit knowledge, the second one relies on fostering the knowledge exchange between individuals. The right strategy typically comprises aspects of both but often puts an emphasis on one of the two. Which of the two is equipped with more resources depends on the relationship between the KM initiative's strategy and business strategy, the affected parts of the company, the individual units of the company, and the processes that are to be supported. Generally speaking, the more creative and more non-standard and complex processes are the more likely a communication strategy is favorable.

4.2.4 Critical Success Factor 2 – Management Leadership and Support

An organization is essentially a collection of individuals who, if left unattended, will pursue largely individual goals (Chourides et al. 2003). The task of senior managers is to provide a *common perspective* and to get people to *work collectively* for common i.e. the organization's goals. Making KM a top priority on the organization's agenda and clearly articulating how it contributes to achieving organizational objectives, helps the organization to be successful, and, thus, is a crucial activity. This is why lack of commitment from an organization's top management can be a serious barrier to successfully implementing a KM initiative (Holsapple, Joshi 2000). This is true for any business initiative with impact on the entire organization (Hasanali 2002, p. 1; Al-Mabrouk 2006, p. 2) and especially applies to initiatives that more fundamentally change the way the company does its business, i.e. in *change management projects* (Davenport et al. 1998, p. 50; Holsapple, Joshi 2000, p. 241; Mårtensson, Maria 2000; Hasanali, APQC 2004, p. 58).

Top management needs to convince the organization's employees that sharing knowledge is necessary and ideally act as *role models* that exhibit the desired behavior Employees need to understand how KM will improve their personal knowledge and work environment in addition to their own potential for success (Davenport et al. 1998, p. 54; Du Plessis 2007, p. 95). To support this, top management must provide the company's employees with a common vision and articulate how they will be affected by new procedures, practices, and technology. If communicated right, the employees will adopt this vision and believe in its contribution, which increases the chance for a KM initiative to enjoy sustainable success (Koenig 2004). Next to communicating a new KM initiative, top management also has to track the progress of the new initiative and evaluate its real contribution (see also 4.2.7), support its adoption by changing existing or introducing new incentives (see also 4.2.11) and to make sure that at all times the necessary resources are available (see also 4.2.14) (O'Dell, Grayson 1998; Holsapple, Joshi 2000).

Summarizing the discussion, top management needs to communicate the value of the KM initiative and get the employees' "buy-in". Furthermore, top management needs to convey how the KM initiative will affect each employee in terms of changes to their processes, technological support, and incentive structure. Finally, top management must act as role model and exhibit the desired behavior.

4.2.5 Critical Success Factor 3 – Organizational Culture

Organizational culture can be described as *shared values, behavioral norms,* and *social customs* and their consequence: the characteristic way in which individuals perform their work and interact in organizations. It is the collective programming of the human mind that distinguishes the members of one human group from those of another group. As such it tends to be deeply rooted in people's hearts and minds, and cannot be changed easily (Hofstede 1980).

The organizational culture also acts as mediator for other success factors such as IT and its adoption (see also 4.2.10) and the way in which the new KM initiative can be managed (see 4.2.4) and thus determines their effects on the overall success of the initiative (Mårtensson, Maria 2000). In addition, it influences which kinds of KM initiatives are possible at all, as it is impossible to force a new concept into an existing culture that do not match (Davenport et al. 1998; Schütt 2003, p. 453).

While a knowledge-friendly culture is often regarded as one of the CSFs for KM initiatives, it is difficult and very time-intensive to establish, in case it does not already exist (Chase 1997; Davenport et al. 1998; Wong 2005). In an ideal "KM culture" new ideas and insights are highly valued and their creation, sharing and application is fostered (Wong 2005). To achieve a suitable organizational culture, some of its aspects have to be taken into account.

Collaboration between individual employees and different groups of employees is essential and a prerequisite for knowledge transfer. This is because knowledge transfer in general, and especially when it is about tacit knowledge, requires people to interact, exchange ideas, and share their experience with one another (Goh 2002). Another aspect is an organizational culture's ability to stimulate and instill an *open-minded attitude* towards expertise and knowledge developed outside of the own field of application, i.e. outside of the own departments or company boundary. This avoids a conscious or unconscious silo behavior, also known as the "not-invented-here" syndrome (O'Dell, Grayson 1998).

Another fundamental aspect of a knowledge-friendly culture is *trust* in the company's regulations to not, directly or indirectly, punish knowledge sharing (Stonehouse, Pemberton 1999; Helm et al. 2007). In some companies, employees are often reluctant to make their knowledge available and transfer it to others in fear of losing their job or value to the company (Davenport et al. 1998). Organizations can build up trust by punishing knowledge hoarding habits and by providing appropriate incentives for knowledge sharing (see also 4.2.11). Another trait of a trust-conveying culture is its *tolerance for making mistakes*. Making reasonable mistakes can be regarded as an investment process in individuals because it can be a source of learning and may foster the creation of new knowledge (Wong 2005). Consequently, the sharing of failures, mistakes, and lessons learned should be tolerated and supported. In this light, trust can also be increased by allowing and supporting employees to question existing practice and by encouraging them to explore new solutions and approaches (Davenport et al. 1998; Stonehouse, Pemberton 1999; Helm et al. 2007).

Summarizing the discussion, the organizational culture is hard to change and determines which KM initiatives are possible and how they need to be tailored, e.g. in terms of management support, incentive structure and technological support. A knowledge-friendly culture fosters knowledge exchange among individuals across departments, incentifies this behavior and punishes knowledge silos. It further supports collaboration, intensifies the employees' trust, that a knowledge sharing behavior is not disadvantageous and that adapting existing practice as well as experimentation to a certain degree is accepted as desirable.

4.2.6 Critical Success Factor 4 – Organizational Infrastructure and Roles

If clearly communicated *responsibilities* for KM are missing, the success of any KM initiative is at stake (Davenport, Prusak 2000b, p. 122). This is why it is imperative for successful KM initiatives that organizations create proper *infrastructures* and appropriate *roles*. Establishing lasting responsibilities and ownership in addition to making departments and employees accountable, ensures the realization of KM initiatives and planned campaigns (Hasanali 2002, p. 3; Wong, Aspinwall 2005, p. 271; Lehner 2008, p. 286).

A KM initiative will only be successful if the ones that also benefit from it, i.e. the knowledge workers, support it. Davenport and Prusak underline this by stating "[...] the most successful organizations are those in which KM is part of everyone's job" (Davenport, Prusak 2000b, p. 107). On the other hand, it is unlikely that all strategic, tactical and operative tasks that come with a new KM initiative can be put "on top" of the existing work load, which is why often new positions specifically for the support of the KM initiative are created or established position's responsibilities are shifted. Several authors have addressed the topic of knowledge roles within an organization and propose their own frameworks for positions they feel are required (Nonaka, Takeuchi 1997; Davenport, Prusak 2000b; Helm et al. 2007). In the following, the characterization provided by Davenport and Prusak (2000b, pp. 108–122) is demonstrated, since it is similar to those of the other authors and also in many regards the most complete. KM activities and responsibilities can be distributed to the following four roles:

- **Knowledge-Oriented Personnel/All Employees:** While specialists are critical to the success of KM, even more important are the activities and attitudes of those who do the operational business of a company. In their day-to-day activities they create, capture, share, search, and apply their personal as well as the common organizational knowledge. In this sense, KM must be part of everyone's job and can be considered part of knowledge-oriented personnel's job.
- **Knowledge Management Workers:** This role includes intermediaries, correctors, editors, group operators and moderators, coaches, and technical facilitators who perform many of the operative tasks and thus offer support to the previous group. They elicit, author, index, structure, categorize, and package knowledge. Technical facilitators additionally configure KMSs to (semi-)automatically categorize, abstract, filter, and disseminate explicitly stored knowledge in directories and document repositories (Soliman, Spooner 2000, pp. 61–62).
- **Managers of KM Projects:** Managers of KM projects are the ones responsible for defined parts of an entire KM initiative They are responsible to develop

objectives and goals within their scope, manage the assigned personnel and other stakeholders, evaluate the benefit of the project and possibly also act as project champion (Davenport, Prusak 2000b, p. 112).

- **Chief Knowledge Officer (CKO):** The most senior role in KM initiatives is assumed by the CKO and is also often regarded as the most crucial and important one (Liebowitz 1999; Davenport, Prusak 2000b; Lehner 2008). CKOs take the lead of parts or the entire KM initiative in an organization, and are responsible for making executive decisions. The appointment of a CKO is usually done with the aim of creating a company-wide integrated perspective on KM. They advocate the importance of knowledge and learning, support any required initiatives and long term cultural, behavioral, and technological changes (see also 4.2.4), oversee internal and external relationships on technological or interpersonal level, evaluate the effectiveness of the KM initiative (see also 4.2.7) and design and implement a KM strategy (see also 4.2.3). But most importantly, they are responsible to establish a knowledge culture, create a KM organizational infrastructure, and to make it all pay off economically (Davenport, Prusak 2000b; Kreibich 2003).

Next to identifying or defining KM roles, a company needs to define the *governance structures* or more specifically how to incorporate KM tasks and roles such as the CKO into the organizational setup. Several organizations have positioned the CKO within their IS/IT, HCM, or Business Management departments (Liebowitz 1999; Davenport, Prusak 2000b). Liebowitz (1999) is generally of the opinion that the CKO should be a member of corporate management while Bullinger et al. (1997) identified three typical forms for successfully integrating KM within the overall organizational structure: In centralistic companies a global KM department is governed by a single CKO. In a shared service scenario, it is likewise but KM support and services are only provided on specifically generated demand by other departments and business areas of the organization. A common global CKO coordinator usually supervises several “local” CKOs and KM departments in decentralized companies. The kind of KM organization suitable for an organization largely depends on its business model and is strongly influenced by its existing infrastructure.

Smaller companies often need to take a smaller scale approach, since they have to consider resource allocations more carefully (Wong 2005). Companies that do not have a dedicated knowledge manager and department usually carry out single KM projects supervised by their corporate management or technological units. In these cases the required responsibilities and decision ownership is placed within the realms of those departments and distributed to their available workforce (Lehner 2008, p. 287).

Summarizing the discussion, the additional tasks for a KM initiative need to be supported by all employees but distinct roles can be identified. Adopting Davenport and Prusak's categorization the general employee, knowledge management personnel, KM project leader, and chief knowledge officer can be distinguished. Next to identifying those roles and their tasks, the organizational governance setup needs to reflect the KM focus. Depending on the company setup and its business, there can be no specific organizational element, e.g. which is often the case in smaller companies, or the responsible department and its head, the CKO, can be placed at centralized or decentralized places within the company.

4.2.7 Critical Success Factor 5 – Measurement

Failing to institutionalize appropriate measurement activities increases the likelihood for KM initiatives to not reach their desired goals and hence to suffer the risk of being labeled unsuccessful (North 2005; Wong 2005). Measurement in this context is a *data collection system* that generates and holds useful information and insights about a particular situation or activity (Wong 2005). Companies need a measurement system in order to understand which goals are realistic, to ensure envisioned objectives are being attained, to track the progress of projects and initiatives, and to determine their benefits and effectiveness. Measuring KM "[…] essentially provides a basis for enterprises to evaluate, compare, control and improve the performance […]"(Wong 2005) of their current KM projects as well as generate lessons learned for possible future initiatives (Ahmed et al. 1999; Arora 2002). In addition, KM measurement is required to *demonstrate added business value* and advantages of a KM project to management and other stakeholders of the company. Without provision of this evidence, leadership support and belief will dwindle and required resources may not be granted (Ahmed et al. 1999; Probst et al. 2000).

KM measurement and the difficulties and problems associated with its activities have been extensively studied and discussed (Probst et al. 2000; Koenig, Srikantaiah 2004; North 2005; Probst et al. 2006; Lehner 2009). In the following, two rather common approaches are described, while for the others the reader is referred to the mentioned literature.

One possibility for measuring KM initiatives' success is *internal or external benchmarking* (Skyrme, Amidon 1997; O'Dell, Grayson 1998; Lehner 2009), which is a KM measuring technique that focuses on measuring and comparing performance in order to identify best practices and determine realistic and competitive KM project goals. A requirement for taking part in a benchmarking activity is that comparative data of organizational departments and external competitors is available. Without such type of data, the basis for comparisons does not exist and benchmarking cannot be performed.

Another way to measure KM success is using *performance evaluations* that focus on measuring the impact and contribution of KM initiatives and the process improvements they offer to the company. The measurement entities can be divided in two categories: *activity-related* and *outcome-related* (Koenig 2004). Activity measures are concerned with measuring knowledge-related actions performed by the organization's employees often using direct measures such as hits on a site or document, number of registered users, number of discussions in a forum or database size of a repository. These measures, however, cannot directly indicate if employees save time and money or if they are able to improve the quality of products and services based on KM efforts. This is why Fahey and Prusak (1998) see (a too strong focus on) activity-related measures as one of the "deadliest sins in Knowledge Management" Nevertheless, direct activity measures can be beneficial to determine the interest and progress of a KM initiative and, therefore, can be beneficial for the KM initiative.

Outcome-related measures in contrast are focused on identifying a change in e.g. sales or margins, hours spent preparing proposals, or cycle time to respond knowledgably to a customer request. Arguably, these types of metrics are a lot more difficult to track (Koenig, Srikantaiah 2004). Users can be asked to grade KM processes and assign a value to the various uses they make of the KMS and new procedures. This grading can be based either on an importance scale or as minutes and hours saved during work. Inquiry forms could for instance include questions on the number of answers found in the knowledge base or the number of answers received from experts found through the KMS. Adding rating criteria to the content and expertise accessible through the KMSs further allows a ratio calculation and tracking of provided useful knowledge as an outcome measure. These techniques are manageable in effort and can be applied in early prototype scenarios to create an insight on possible improvements for a currently running KM initiative. The scope and type of the users can be gradually enlarged as projects become more mature in order to receive more widely ranging and representative results (Laudon, Laudon 2006).

Summarizing the discussion, measuring the effects of KM initiatives allows the responsible managers to further promote the KM initiative and to acquire more resources for it. There are a large number of approaches to measure the contribution of a KM initiative; however, each has its drawbacks. One exemplary approach is benchmarking which compares the processes and their results within an organization company or across organizations within a benchmarking group. Another exemplary approach is performance evaluation which can either focus on activity-related measures, i.e. how intensively KM tools and processes are used, or on output-related measures, i.e. how large the contribution to the business metrics such as sales figures are.

4.2.8 Critical Success Factor 6 – Knowledge Structures and Ontologies

Knowledge and its categorization can be fuzzy and often is different for each person, which is why every individual has his own way of structuring what he knows. On the other hand, there should be *generally accepted structures* to facilitate the transfer of knowledge. Finding the right balance and supporting a common understanding is critical for the success of KM initiatives. This is why a common language, i.e. a vocabulary understandable to all affected employees, plays a key role in enabling an organization's members to communicate and collaborate with each other. Consequently, for KM initiatives to be successful it is vital that employees share a common understanding of the terms that are used within the organization and are "on the same page" when confronted with knowledge-intensive activities during their daily routines. Davenport and Prusak (2000b) offer support for this argument. They found that companies who had conducted successful KM projects were also the ones where employees understood the concepts of knowledge and KM well.

Considering the management of explicit knowledge, it is the goal to collect the information of the organization within a suitable structure, and hence adding the necessary context in form of e.g. keywords or tags. Organizations that manage explicit knowledge successfully in their initiatives use the established vocabularies as a foundation to create specific knowledge categories and navigation mechanisms for their organizational knowledge base (Liebowitz 1999; Laudon, Laudon 2006; Lehner 2009). Often, the organization's *knowledge map* is reflected in *ontologies* and *database schemes* that are used in KMS such as knowledge repositories or supportive technology such as search engines. In IS research, an ontology is a formal digital representation of a set of categories and concepts and the relationships between them (Liebowitz 1999). Ontologies form the basis for supporting the knowledge maps and structures of the organization with IT and therefore allow fast and easy access, as well as retrieval of documents and other electronically stored knowledge by computer search engines (O'Dell, Grayson 1998; Liebowitz 1999; Lehner 2009). Using enterprise search applications, knowledge structures can be used to define what can be searched and how the results should best be displayed.

Considering the management of implicit knowledge, it is the goal to transfer knowledge that is not explicable or at least it is not easy to do so. Rather, it is transferred by bringing together the *right* persons that subsequently exchange the knowledge in a tacit way. Having company ontologies in place can also help such approaches by e.g. offering suitable means to identify capabilities in YPS and by offering first hints towards a *common vocabulary* that can be used in the online or offline communication during the process of knowledge exchange.

An additional challenge when dealing with knowledge structures or ontologies lies in keeping them up-to-date. The sum of individual's knowledge and the implicit knowledge inherent to the company that is coded in its processes and codes of conduct is not static but changes over time. Furthermore, the terminology used within the company may change over time, adapting to external changes e.g. new technology that is available or to internal changes e.g.

new products in the company's product portfolio. This is why knowledge structures should reflect the *current usage pattern* of the stakeholders (Davenport et al. 1998).

Summarizing the discussion, categorizing the knowledge within a company helps the knowledge exchange process, because different persons in a company have a reference vocabulary to rely on. The categories are represented in knowledge structures such as ontologies. This meta-information supports storage and retrieval of explicit knowledge by e.g. providing a blueprint for the internal storage format. Implicit knowledge exchange is supported by offering a "reference language" during tacit knowledge transfer and when searching for capabilities in e.g. YPS. Knowledge structures face the challenge of being kept up-to-date as an organization's knowledge, its usage and interpretation change over time.

4.2.9 Critical Success Factor 7 – Processes and Workflows

The development and execution of KM processes is a central task in any KM initiative (Wong 2005). Essentially, KM initiatives try to improve, support, and facilitate knowledge-related activities and processes of an organization. They do so by introducing well-defined workflows and processes for those knowledge-related activities or more precisely by adapting current workflows and processes with KM activities that benefit the organization.

For a KM initiative to be successful, the development and execution of KM processes is fundamental (Wong 2005). If established in a way that fits the organization's needs and its business situation the KM processes ensure an effective and efficient usage of an organization's knowledge (Holsapple, Joshi 2000; Du Plessis 2007). In order to understand which processes an organization should support most, the *interplay of KM processes* needs to be looked at and mapped to the organization's setup. The most widely accepted categorization of KM processes has been discussed as part of the fundamental concepts for this thesis in section 2.2.3.

Tying KM efforts as closely as possible to the value-creating business processes of an organization ensures that knowledge is likely to be used to optimize and build the critical capabilities of the organization (Donoghue et al. 1999). Some like (Abecker et al. 2002) and (Heisig 2005) take this argument even further and reason that successful KM not only requires focusing on improving the core and enabling processes, but that it is essential to *integrate KM activities* into the daily working routines of employees and into workflows of existing KMS. This way, the enactment of the business processes itself is supported and KM activities are not seen as additional burden but as part of the actual business process. As a result, the achievement of an organization's competitive strategy simultaneously improves the organizational knowledge base as a "by-product" (Markus 2001). In the same line of argumentation, Qureshi, Briggs and Hlupic (2006) showed that it is important to embed knowledge flows into the work context to build up less of a barrier for using KMS.

Summarizing the discussion, the development and execution of KM processes is fundamental for organizations to be successful in KM which is why the processes need to fit to the organization's requirements. There are some suggestions for formalizations of KM processes of which the nine-category Probst/Raub/Romhardt model previously discussed in section 2.2.3 is most prominent in the German-speaking KM community. Additionally it is essential to embed the KM activities suitably into the operational processes. Ideally, the user does the (possibly adapted) business process activities and the KM activities are "by-products" without the user feeling an extra burden but rather additional support.

4.2.10 Critical Success Factor 8 – Information Technology and Knowledge Management Systems

Some KM initiatives are completely "non-digitalized", i.e. they do not need any IT support at all. For example, the setup of informal meetings of experts every month to exchange "war stories" might for a certain company pose a very useful measure to spread the implicit knowledge of its experts and it does not rely on any IT support. However, for many KM initiatives IT support is critical for success because it functions as an enabler for solving some of the inherent challenges companies face when dealing with knowledge and the implementation of KM initiatives (APQC 2000; Wong 2005). The two trends - increasing *information overload* and the *increasing distribution of knowledge* among different potentially distant employees that are expert in their respective fields - fuels the need to employ IT to support KM initiatives. Also, IT and more specifically KMS have been attributed to allow the design and implementation of especially *efficient* and *effective* KM processes and workflows (APQC 2000; Alavi, Leidner 2001; Wong 2005; Al-Mabrouk 2006; Lehner 2009).

All of the nine sources in Table 4-4 acknowledged that proper support of IT, i.e. proper usage of KMS is a critical factor for the success of a KM initiative. However, none of them offered a deeper insight in what renders KMS' contribution to the KM initiative advantageous. To get a deeper insight into the KMS contribution and also because this thesis' suggested KM approach relies on KMS support, the initial literature review is extended in a second iteration where the focus was shifted to determining which aspects of KMS make the KM initiative successful. The results of this second iteration are elaborated on in sub-chapter 4.3.

4.2.11 Critical Success Factor 9 – Motivation

For KM projects to be successful, it is vital that organizations motivate their members to use the newly established infrastructures, processes and systems, to share and apply knowledge, and to contribute to the enhancement of the common organizational knowledge base. Without *provision of appropriate incentives* the whole KM initiative might not be accepted and is at stake to fail (Davenport et al. 1998; Wong 2005; Lehner 2009).

Motivating employees to accept changes of any kind and new KM procedures and systems in particular is especially important at the beginning of the change initiative. When motivated to use the system, the actual usage will increase and the trust in the resource increases. Probst et al. (2006) show the inverse in their *death spiral* of a knowledge resource (see Figure 4-3). They illustrate what happens when a knowledge resource enjoys no trust, i.e. employees are not motivated to use it. When useful content of sufficient quality is not provided in the beginning of initiatives, and maintained at later stages of them, negative cycles can be the result. The scarce amount and lack of quality knowledge resources further reduces the trust of employees in the new system and processes. Subsequently, usage goes down, leadership support dwindles, and further financial investments are not made. This is obviously not beneficial to the quantity and quality of the knowledge resources and thus enforces the negative cycle, which in turn eventually leads the KM initiative to its unsuccessful end.

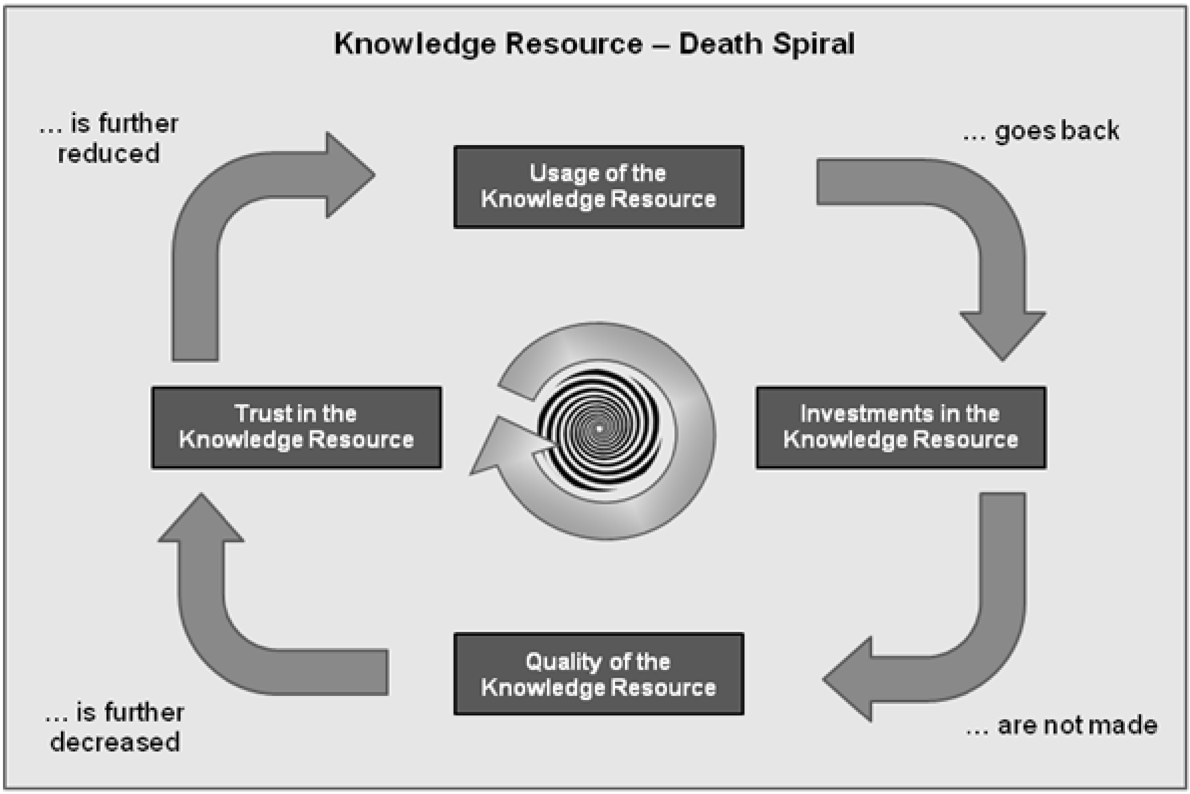


Figure 4-3 ***Probst et al.'s Death Spiral that leads to failure of KM initiatives***
Source:(Probst et al. 2006, p. 207)

In line with their argumentation, Probst et al. (2006) argue that it is especially important to provide a *critical mass of useful knowledge resources* of sufficient quality during the launch of KM initiatives This way, a number of early adopters can be attracted who will support and sustain the system.

One way of achieving the necessary acceptance and hence, also usage and trust in new knowledge-related systems and processes is to clearly highlight the value they supply by demonstrating how they are beneficial to the daily work of employees. Members of the organization want to understand how their work becomes faster, richer, and more rewarding and how the organizational KM efforts contribute to achieving their project and work objectives more easily and efficiently (O'Dell, Grayson 1998; Koenig, Srikantaiah 2004). Organizations trying to conduct successful KM initiatives, therefore, have to focus on supplying real improvements and benefits and demonstrate their added value at early stages of their KM initiatives. When KM initiatives are for instance based on KMS, it is essential for the success of the project that the designed systems are not only effective but also offer “great content” and useful knowledge resources from the start (Koenig, Srikantaiah 2004, p. 64; Krcmar 2010, p. 652).

Next to motivating a company’s employees to use a KMS in the role of a consumer, KM initiatives also strive to make the knowledge of an individual employee available for others in the company. In this case, the individual employee needs to share his knowledge and for this to happen, must be motivated to do so. Creating and sharing knowledge by e.g. writing documents and entering them into knowledge repositories for general use is time consuming, can require substantial effort, and is often considered an extra burden. The same holds for KM approaches that build upon transfer of tacit knowledge such as story telling or physical meetings of experts that are also time consuming and require resources.

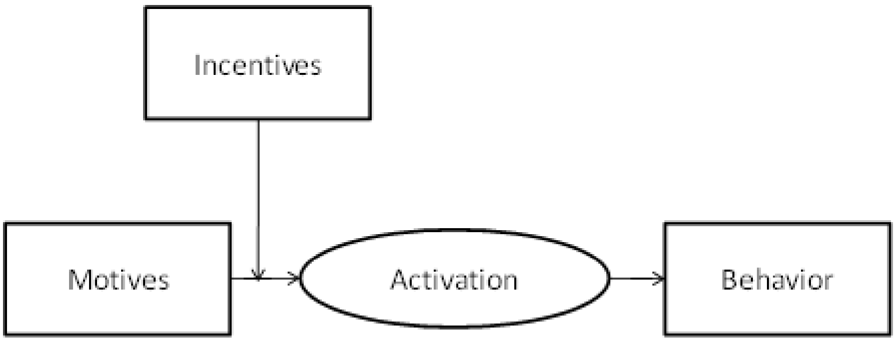


Figure 4-4 ***Rosenstiel's MIAB model describing motivation***
Source: Adapted from (von Rosenstiel 2007)

Motivation to share knowledge depends on the individual’s *motives*. Psychology identifies a motive as an individual’s *psychological disposition* (Lakhani, von Hippel 2003; Heckhausen 2006). It represents an individual’s values and stays relatively unchanged over time. During the socialization process of an individual a set of motives is developed (Lakhani, von Hippel 2003). In a certain situational context, a motive can be activated and subsequently leads to a behavior of the individual. The term *activation* refers to the individual’s response to a *perceived stimulus*, either internal such as a personal desire or external such as monetary compensation. This interplay is referred to as *motivation* (Schneider, Schmalt 2000). The described interactions are part of Rosenstiel's MIAB (Motive – Incentive – Activation – Behavior) model (von Rosenstiel 2007) of an individual's motivation as illustrated in Figure 4-4.

Motivation can be distinguished according to different aspects, but the most often used and hence most important in the context of KM is the distinction of *intrinsic motivation* and *extrinsic motivation* (Nerdinger 1995, 2003). Intrinsic motivation results from the mere execution of an activity. If an individual "enjoys" doing an activity, because he is interested, it makes him content, the activity is perceived as fun, the individual does the activity due to intrinsic motivation. Intrinsic motivation in the context of KM is more likely if employees are involved in the design phases of KM initiatives and identify with the final results, new processes and KMS (Krönig 2001; North, Varlese 2006). Any other motivation that does not result from the execution of the activity alone, is referred to as extrinsic motivation. If an individual does an activity to earn money or to increase his reputation, this actor is extrinsically motivated.

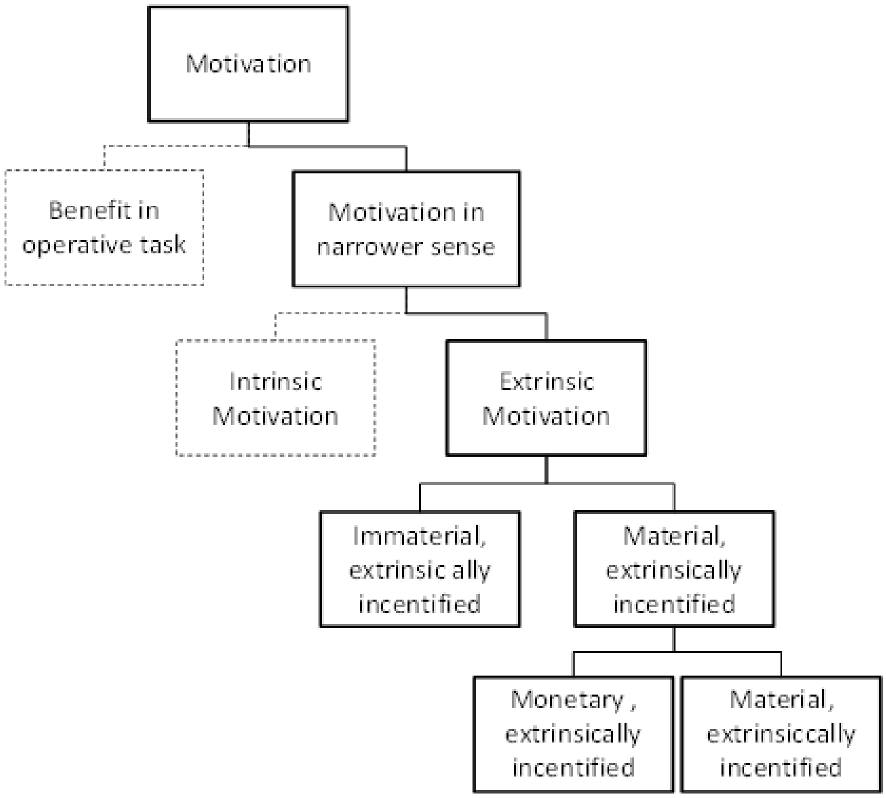


Figure 4-5 ***Different forms of motivation***
Source: Own illustration

In the following, different possibilities to motivate employees will be discussed, referring to Figure 4-5 for illustration.

One way to "motivate" employees has been shortly addressed in the discussion before. When KM processes are seamlessly integrated into an employee's tasks within a business process, the additional effort is hidden in the operational process. This way the employee attends his everyday efforts and simultaneously improves the organizational knowledge base as a “by-product” of the employee’s daily work (O'Dell, Grayson 1998; Markus 2001). He is indirectly motivated to conduct KM tasks by being motivated to do his daily work, as indicated by the dashed lines of the top left rectangle in Figure 4-5.

Properly motivating employees needs additional measures in order to achieve a more complete contribution from most or even the entire workforce. Therefore, it is necessary to understand what motivates employees to share their knowledge. Adopting the results of (Krcmar 2010) and (Davenport, Prusak 2000b), members of a company are most likely to contribute and share knowledge in the following four cases:

- When they want to express an *altruistic attitude* and like to contribute to the common wellbeing (intrinsic motivation)
- When they are *internally motivated* and convinced that sharing is needed to guarantee the success of the enterprise (intrinsic motivation)
- When they can increase their *own reputation* and have the chance to earn the respect of superiors and peers (extrinsic motivation)
- When they can expect *concrete immediate or future direct rewards* (extrinsic motivation)

In a KM initiative there should be a carefully determined balance between building upon intrinsic motivation and supporting the motivation with extrinsic (explicit) rewards such as compensation systems (Krönig 2001). Intrinsic motivation cannot be influenced as such, which is why it is drawn in dashed lines in Figure 4-5. A first step in motivating employees often is to proclaim the individual recognition of an individual’s good KM conduct and celebrate exceptional contributions made by this individual. Employees like being acknowledged, respected by superiors and admired by peers. As Figure 4-5 shows, other *explicit* rewards can be of *immaterial* or *material* nature and the later can be either be *financial* or *material* benefits (O'Dell, Grayson 1998; Davenport et al. 1998; Lehner 2009). Immaterial explicit rewards extend recognition and acknowledgement of excellent contributions from employees by e.g. also considering them for promotions (Koenig, Srikantaiah 2004). Further examples include providing extra free time and vacations or offering personal growth opportunities, such as educational seminars and expertise development trainings (North, Varlese 2006). The later example could offer additional benefit to the company, because any expertise and educational program that expands the skill set of employees can in turn be beneficial for the organizational knowledge base as a whole.

Materialistic extrinsic rewards can take many different forms. Some organizations handed out incentives as mundane as chocolate bars and mouse pads, others have been more “generous” and offer substantial rewards including frequent flyer miles, company-paid trips to recreation resorts or new laptop computers (Davenport et al. 1998).

Financial explicit rewards are usually tied to the performance evaluation and compensation systems of the company. As part of their yearly performance reviews, companies like most consultancies partially evaluate their employees on the quantity and quality of the knowledge they contribute to Knowledge Repositories or CoPs and on how they apply existing knowledge resources to solve daily tasks (Davenport et al. 1998; Hansen et al. 1999; Liebowitz 1999). Wong (2005) as well as Yahya and Goh (2002) argue that enterprises should restructure existing *performance evaluation systems* by not just focusing on individual contributions and goals. Establishing goals that individuals can influence but not achieve on their own and rewarding employees with a focus on team or group performance can further support teamwork, collaboration and socialization can facilitate knowledge transfer and exchange.

North and Varlese (2006) argue that motivation can additionally be increased by adopting creative solutions like *company-wide games* and *competitions* that can be a useful supplement and have a substantial motivational effect.

Finding the right mix of incentives and motivational aids is highly dependent on the unique circumstances of an organization and a key factor for achieving successful KM projects. Nevertheless, there is no common agreement on how motivation can best be increased. Some claim, that extrinsic motivation is an inadequate measure entirely. Schütt (2003, p. 453) for example, indicates that a company that formally relied on an extrinsic materialistic reward stopped using this motivational vehicle as it did not work as expected.

Another way to increase the motivation to share is to empower KM workers that act as *facilitators*. They are hired to make sure that everyone contributes to the company's knowledge base. This approach, however, only works as indirect motivation, because employees are not motivated by the mere presence of a facilitator but by the possible incentives and punishment that this facilitator may use. Additionally the use of specific KM personnel should be treated with care as KM efforts are most successful when they are part of everyone's job (Davenport, Prusak 2000b) and less so when they are the job of a few that push the effort onto the many without the many benefitting (directly) from the effort.

Summarizing the discussion, motivation of employees is a key aspect as it is the employees that either live or abolish the KM initiative. This is especially true in the beginning of a KM initiative where the risk of failure is larger as the new ways of performing the tasks need to be justified with some benefit. If this cannot be obtained the death spiral of the KM initiative may lead to an unsuccessful end. Employees are indirectly motivated to do KM tasks if they are embedded and hence hidden in their operational tasks or if they show obvious benefit. The contribution of knowledge for the success of the KM project is harder to achieve as it almost always includes extra effort. Employees can be intrinsically motivated to contribute, when they feel that supporting the KM initiative "is the right thing" to do or when they live up to an altruistic behavior. Employees' motivation can be supported with external incentives that can be material or immaterial. While material incentives can be monetary or non-monetary, e.g. company-paid vacations, immaterial incentives can comprise higher reputation. The right balance between relying on intrinsic motivation and sparking the engagement with extrinsic motivation is crucial and arguments go as far as indicating that extrinsic incentives are never a good measure. In the initiation phase, however, extrinsic motivation can help bridge the phase in which high value content is necessary to carry the KM approach afterwards (Oldigs-Kerber 2007, p. 69).

4.2.12 Critical Success Factor 10 – Training

Training employees and offering tailored educational programs is regarded as an important measure to ensure the success of a KM initiative (APQC 2000; Goh 2002; Chourides et al. 2003; Wong 2005). Training can be interpreted in a broader way to also cover informing the employees about the changes and how they affect the employees, i.e. it covers *marketing the change* to the employees. KM initiatives change the way employees do their tasks and how the company as a whole operates and hence can be a substantial change to the well-established status quo of an employee. Similar to other change management projects, KM initiatives face the challenge to address employees' *barriers to change*. Reiß (1997, p. 17) identified four typical personal barriers that are depicted in Figure 4-6.

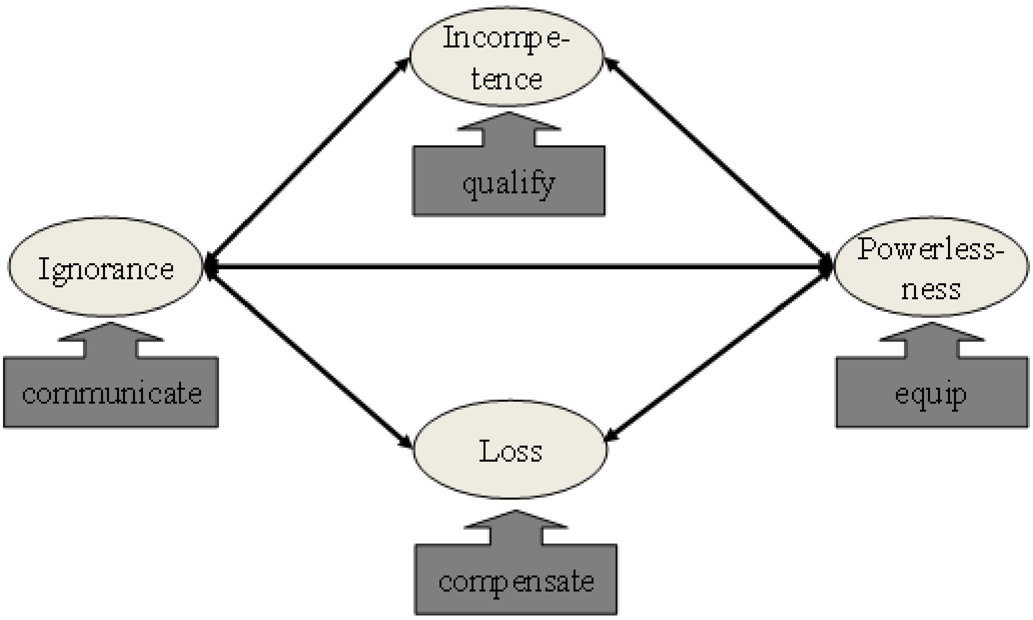


Figure 4-6 **_Barriers to change and how to address them_**
Source: Adapted from (Reiß 1997, p. 17)

If not adequately addressed theses barriers may hinder the employee's acceptance and hence doom the KM initiative to fail. This is why KM initiatives need to be marketed well to address these barriers. Training programs serve this purpose and support the introduction and sustained adoption of a KM initiative. Table 4-5 explains the possible barriers of Figure 4-6 in more detail and suggests how they can be addressed through means of KM-related trainings.

Barrier	Interpretation	How to address
Ignorance	Not knowing about the change	Inform about new tasks and processes, the reason of change and how the individual will benefit
Incompetence	Not knowing how to do the new process; lacking ability to assume new roles and tasks	Educate and train on new process and tools
Loss	Fear of personal loss and as a consequence no motivation to participate	Explain the compensation of additional efforts, reward of expected behavior, inform about personal and organization-wide benefit
Powerlessness	Not being empowered in the new structures; remainders of "old" social / organizational power	Inform about new roles, duties and new responsibilities as well as escalation mechanisms

Table 4-5 ***Barriers to KM initiatives and how to address them in the context of KM trainings***
Source: Own Illustration

Ignorance in this context means that employees simply do not know about the new way of doing knowledge-related tasks. Without knowing the new processes and tasks, they, of course, cannot live them. In trainings the new tasks are marketed, the reason of change is explained and the individual learns how he will personally benefit. *Incompetence* refers to the employees' inability to mentally comprehend the new ways of doing their tasks. The necessary know-how for the new processes or KMS should be conveyed so that the employees know how the new process is to be enacted and are able to use the tools associated with the new process. *Loss* refers to the employees' fear of degradation in value to the company and within their social group and hence also loss of power. The "old" ways of doing have created a structure within the group in which every individual knows how he contributes to the group's well-being. In the "new" ways of doing the personal spot in the emerging structure is not clear yet and it is a typical behavior to fear that it might be worse. Consequently, trainings should emphasize how each individual can create benefits for the organization by adopting the new processes and also how the individual will personally benefit. Finally, *Powerlessness* refers to the initial inability of individuals to adapt to new governance and power structures, especially in established groups that have strong and long-established power and delegation mechanisms. Those existing structures are changed due to the new KM processes, which is why the empowerment that comes with the new processes needs to be communicated and escalation mechanisms must be made clear to break up the established governance structures.

Following the previous argumentation, it becomes apparent why KM initiatives are at risk to be unsuccessful if they fail to market the purpose of KM efforts. Designing and introducing

KM initiatives and the accompanying tools, is only advantageous if the company's employees understand their *purpose*, *identify* with the initiative, its purpose and goals, have sufficient *know-how* to use them, understand their *personal benefit* and their *new roles* (Koenig, Srikantaiah 2004). Consequently, training programs should be regarded as investments in an organization's people (Chourides et al. 2003) and this way in the organization's value in the form of intellectual property.

Relating to the other CSFs and the possible barriers mentioned above, KM-related trainings can focus on the following:

- **Significance of KM initiative**: A focus lies on creating a *common vision* and *goals* for the company's employees. It must be made apparent to all employees how the company and each individual will benefit. Possibly a feeling of urgency to adopt the KM initiative is created in the process. Trainings with this focus address the *ignorance barrier*.
- **Defining a common language:** Having a *shared common language* allows employees to work and collaborate more efficiently, to reduce misunderstandings especially between employees of different departments, and fosters a feeling of belonging to the same entity. Trainings with this focus try to address the *ignorance barrier* as well as the *incompetence barrier*.
- **KM Strategy and KM approaches:** Increasing the employees' understanding when to use which channels of knowledge transfer in dependence to the kind of knowledge sought increases the KM initiatives efficiency. Additionally, communicating and explaining how and why certain KM strategies were chosen for the corporation in regard to its competitive strategy supports *procedural transparency* (Folger, Cropanzano 1998). This in turn leads to employees that exhibit more *motivation* (Nerdinger 2004). Trainings with this focus may address the *ignorance barrier* and the *powerlessness barrier*.
- **KM Governance:** As with other projects, it must be obvious who is the *project sponsor* and/or *project champion*, i.e. employees must understand who "owns" the KM initiative and is responsible for it. However, after its introduction, it is additionally necessary for each employee to understand which specific positions have been created or are part of KM initiatives, what they are *personally responsible* and *accountable* for, and who to address their requests to (Koenig, Srikantaiah 2004, p. 68). Trainings with this focus address the *powerlessness barrier*.
- **Integration of KM and business processes:** An organization's employees best adopt KM initiatives when their *benefit* is obvious. This is why the focus in trainings can be put on explaining the new processes and demonstrating how they make

employees’ life easier. Trainings with this focus address the *loss barrier* and the *ignorance barrier*.

- **KMS usage:** Even when new software, such as KMS, should ideally be self-explanatory, employees should be trained and educated in the use of the new KMS. This measure ensures that employees understand how the systems can be configured, what functionality they offer, and how their full capabilities can be utilized. Without provision of such trainings even the best systems are at risk of not being accepted and can become graveyards of information and knowledge (Koenig, Srikantaiah 2004; Wong 2005). Trainings with this focus address the *ignorance barrier* and the *incompetence barrier*.
- **KM rewards and performance evaluation:** When relying on *extrinsic incentivation* (see 4.2.11), employees should be informed about the rewards and performance evaluation that are introduced along with the new processes and tasks. This way these motivation-increasing measures can unfold their full potential. Trainings with this focus address the *loss barrier*.

While trainings are pivotal in the implementation phase of a KM initiative to market the added value and the necessary changes, ideally, KM initiatives should also offer the potential for *continuous learning* and *knowledge development* at all levels within the company. This can be achieved by offering an environment that encourages individuals to ask questions and by challenging them to learn according to their unique capabilities (Skyrme, Amidon 1997).

Summarizing the discussion, KM trainings allow organizations to support the introduction and sustained involvement in a new KM initiative. They should be tailored to address the common barriers to change: *ignorance*, *incompetence*, *powerlessness* and *loss* (Reiß 1997). Trainings should serve the purpose of informing the employees about the new initiative, educate them to be able to adopt it and use the associated tools, inform them about new responsibilities and organizational structures, and explain the reward they get from adoption. It should be made obvious how the KM initiative contributions to the benefit of the individual and the organization. Trainings can have different focus areas in terms of their content that relate to the mentioned barriers and that are tailored to moderate the other CSFs’ impact. In a different notion, training should also comprise the possibility for *continuous learning* in the organization and encourage adaptation of the KM initiative to respond to changing demands.

4.2.13 Critical Success Factor 11 – Human Capital Management

“Managing knowledge is managing people; managing people is managing knowledge” (Davenport, Völpel 2001). While some parts of an organization’s knowledge may be stored in tangible ways in databases or patents, it are essentially the employees of a company that acquire, identify, develop and apply the organizational knowledge in order to create value and achieve business objectives. This is why CKOs and knowledge managers who supervise and conduct KM initiatives need to integrate their efforts to the management of this resource. They need to be aware of the possible benefits of effective, KM-oriented HCM and the fact that it has a strong influence on the sustainability and success of their KM initiatives (Chourides et al. 2003; Wong 2005).

The importance of the social dimension and the vital role of HCM in KM have been vividly discussed (Robertson, O'Malley Hammersley 2000; Soliman, Spooner 2000; Zeleny 2005). Three areas of HCM have been mentioned frequently to show a strong interaction with KM aspects and will be addressed in the following: *recruitment* of new employees, the *professional development* of the company's workforce and the *retention* of employees (see Figure 4-7).

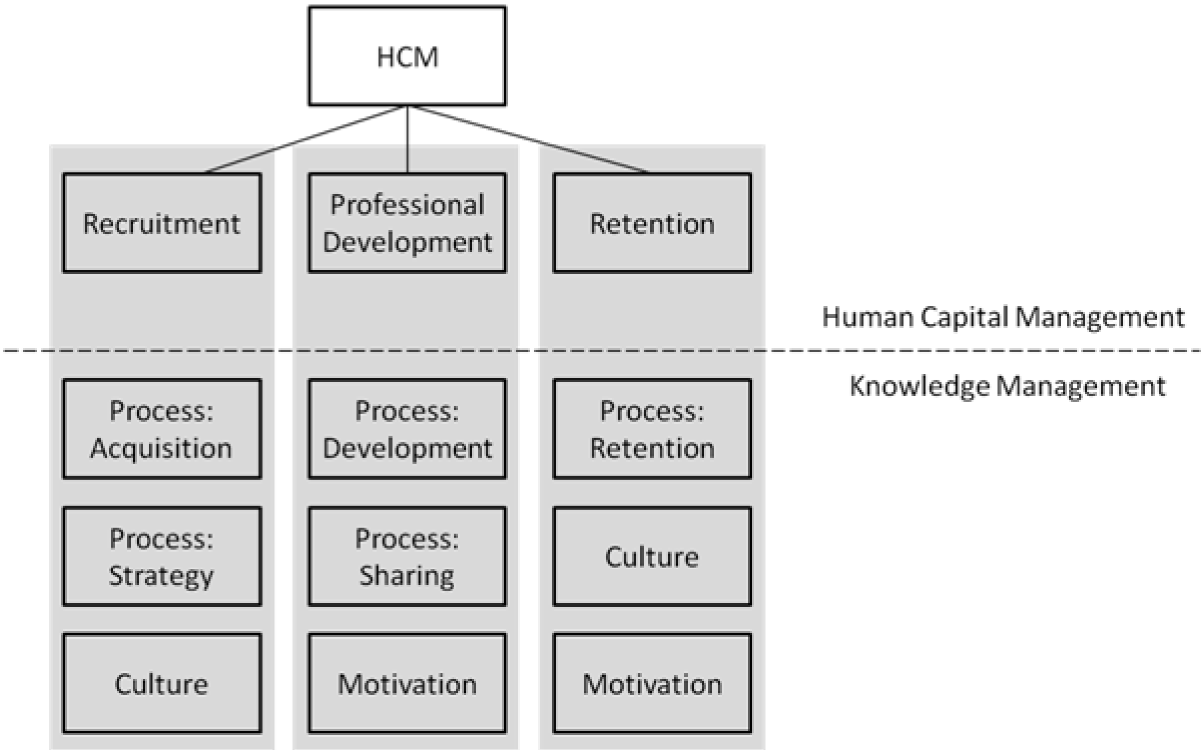


Figure 4-7 ***Relationship between the three HCM areas most related to KM aspects***
Source: Own Illustration

Some authors argue that people are the most important asset of an organization (Chourides et al. 2003). Certainly, this statement has to be interpreted with respect to the business of the organization and is not true for all. However, the more knowledge-intensive the organization's business is, the more it will be the case that the organization's employees and their knowledge and abilities become the determining factor for success. Effective expansion of the workforce by recruiting new employees is thus a critical HCM activity in regard to the success of the organization in general and of KM initiatives in particular. The *recruitment process* regulates the expansion of company's taskforce according to its needs. On the one hand, it can be used to acquire new external knowledge and competencies for the company in line with the KM process *Knowledge Acquisition* (see 4.2.9). In terms of the KM process *Knowledge Strategy* (see 4.2.9), it also represents the main means through which identified strategic knowledge gaps inherent to the current staff can be addressed (Soliman, Spooner 2000). On the other hand, by filtering candidates that apply for the company it is also possible to slowly adapt the *culture* of an organization (see 4.2.5). The cultural fit is at least as important as fitting the skill set of new employees to a job specification because KM thrives on people being "part of it", having the ability to collaborate with each other as well as seeing the inner purpose of sharing their personal knowledge in order to serve the common goal. Therefore, special care needs to be taken in order to assure that new employees fit in with the distinctive ways of how work is performed and conducted within the organization. Hence, companies should choose employees according to necessary and missing skills for the achievement of their core KM related tasks, but also based on cultural fit. Choosing fitting new employees and establishing harmonious and well-working environments and teams is thus often regarded as a fundamental aspect for success in knowledge-related projects (Robertson, O'Malley Hammersley 2000; Goh 2002).

Professional development programs and trainings focus on improving the competencies, skills and knowledge of individual employees of the company. Developing the employees of a company increases their abilities, extends the responsibilities they can assume, and their potential to contribute value to the company. Additionally, however, it can also benefit the organizational knowledge base as a whole. Skilled employees can use their extended competence to improve process quality and to increase efficiency. Since processes are "encoded knowledge" of the company, the common organizational knowledge base benefits. Innovations that are part of the *Knowledge Development* process (see 4.2.9) are, in the same line of argumentation, more likely if the employees are educated well. In addition, if the company is able to support the *Knowledge Distribution* process (see 4.2.9), it is possible to spread the new knowledge of the employee to other employees and by this means to achieve increased benefit. More capable and educated employees consequently can help the enterprise to create valuable services and products and as a result to compete successfully on the market (Robertson, O'Malley Hammersley 2000; Soliman, Spooner 2000; Wong 2005). Additionally, *career development* can be a motivation and can be used as an *incentive mechanism* as well (see 4.2.11). The individual employee can regard the investment into his professional career as positive feedback, that not only motivates, but establishes a higher degree of loyalty to the organization and hence, supports the *retention activities* of the company

Retaining employees is an important goal for HCM and KM initiatives likewise. If an employee leaves, the company not only looses one unit of human labor, but more importantly, the tacit knowledge of this employee is lost as well. While formalized knowledge can be replaced more easily by hiring a new employee with the necessary knowledge, implicit knowledge is not replaceable that easily. It entails an employee's experience of how to apply his knowledge in the context of the organization and meta-knowledge about the company's knowledge sources, i.e. where to find help and advice in either stored format or in co-workers. This way, the shared organizational knowledge is decreased when an employee leaves and, if this employee was an experienced member, limits the ability to engage in *Knowledge Distribution* and *Knowledge Development* processes. This is why *Retention* is also a process in KM (see 4.2.9), where the preservation of knowledge takes center stage. Retaining employees and preventing knowledge loss whenever possible is therefore an important task in KM and HCM (Wong 2005). Recruiting and employee development activities, as discussed above, can also be considered important building blocks for achieving employee retention. When employees are recruited that fit into the company's *culture* (see 4.2.5) and when they are offered desirable development possibilities and career chances it increases the likelihood that they will feel comfortable and identify with the company and thus that they stay with the company (Robertson, O'Malley Hammersley 2000; Zeleny 2005). On the other hand, when employees leave the company and are replaced by new individuals, the *shared believes* of the company may shift, i.e. the culture may change. A fundamental way to retain employees is keeping them happy. If they are happy, employees identify with the company and are less likely to leave the company. Seen from a KM perspective, employees that identify with the company are more eager to pursue the shared goal of the company and are hence more likely to feel that they should contribute to the organization's knowledge base. If they are content with their work, they are *intrinsically motivated* (see 4.2.11). As a result they could be more inclined to help transfer their critical knowledge and make it available to the common knowledge base before they leave the organization (Robertson, O'Malley Hammersley 2000; Zeleny 2005).

Summarizing the discussion, managing employees is a major part of managing knowledge, which is why HCM and KM interact strongly. Three areas of HCM are especially important for KM: *Recruitment*, *professional development* and *retention*. Recruiting is concerned with hiring persons who possess the necessary formal qualifications and this way extend the organization's knowledge base but who also fit into the d the organization's *culture* and help to reach the organization's strategic knowledge goals. *Professional development* is concerned with educating and training employees to acquire new skills and abilities. Better educated employees may share their acquired knowledge, may improve the company's processes, both measures to increase the company's knowledge base, and they are more likely to create innovations, i.e. to develop new knowledge. In addition, investment into professional development can improve an employee's *motivation*. *Retention* is concerned with making employees stay with the company as long as possible. From a KM perspective, this is necessary to not lose the knowledge that employees possess, especially the implicit knowledge, which is not easily replaceable. Employees are more likely to stay when they fit into the organization's culture and, on the other hand, if they do leave and are replaced, may slowly change the culture. A way to keep employees as long as possible is by keeping them content with their work. This way they are also more likely to be motivated to share their knowledge because they identify with the company and its goals.

4.2.14 Critical Success Factor 12 – Resources

Similar to other kinds of initiatives, successful KM initiatives depend upon the amount of resources that the company allocates to them. Available resources govern the possible quantity and quality of the efforts that are directed towards implementing KM within the company (Holsapple, Joshi 2000; Wong 2005). The availability of resources also moderates the influence of the other CSFs. For example, developing a suitable *KM strategy*, setting up *knowledge structures* and designing efficient *KM processes* requires the availability of competent employees, hence human resources. Human resources are also necessary to staff KM-related *roles* and to establish KM departments within the organizational infrastructure of the enterprise. On the other hand, financial resources are necessary for example to implement, configure, and maintain *KMS*. In addition, time is also a critical resource. Companies must give their employees time slots that they can allocate to identify and acquire existing organizational knowledge and best practices and to enable them to share and contribute knowledge to the common organizational knowledge base. In order to have time to spend on KM activities, other operational tasks have to be allocated less time. So, at least in the beginning of an initiative, before the benefit of the KM initiative outweighs its costs, the initiative will likely decrease productivity in terms of time spend for the same output. Time as well as financial resources for designing and performing training activities are also seen as essential (Mårtensson, Maria 2000; Yahya, Goh 2002; Wong 2005).

None of the previously identified CSFs alone will make a KM initiative successful; rather all of them have to be adhered to in concert. This is true also for proper support with resources. Without provision of required resources, maintaining a sufficient degree of perfection considering the other CSFs is hardly possible. Resources, therefore, enable the proper implementation of the other CSFs as shown in Figure 4-8. Someone who has authority over budget – typically someone in a senior management or other leadership position – grants resources to an initiative. This fact is illustrated in the left hand part of Figure 4-8. As illustrated in the upper left part of Figure 4-8, in order to get *leadership support* (see 4.2.4) and consequently *resources*, a KM initiative must show economic benefits and its impact to creating value for the corporation should ideally prove positive in advance of the initiative (Davenport et al. 1998). However, measuring the achieved benefit is often hard and providing somewhat exact financial figures and demonstrating the impact of KM initiatives can be impossible, even after the initiative has been implemented (see 4.2.7). This is why *top management sup-*

port is essential for allocating the necessary resources. Without dependable figures, especially before the initiative and during its launch, the provision of resources for KM initiatives is subject to management's belief that it will work. This is similar to other organizational activities such as marketing campaigns or social events such as company-sponsored celebrations of anniversaries. In those cases, it is often also not possible to determine the direct (financial) impact but it is a "gut feeling" of the responsible senior managers that good marketing will promote sales efforts and that social events will increase employees' satisfaction and identification with the company.

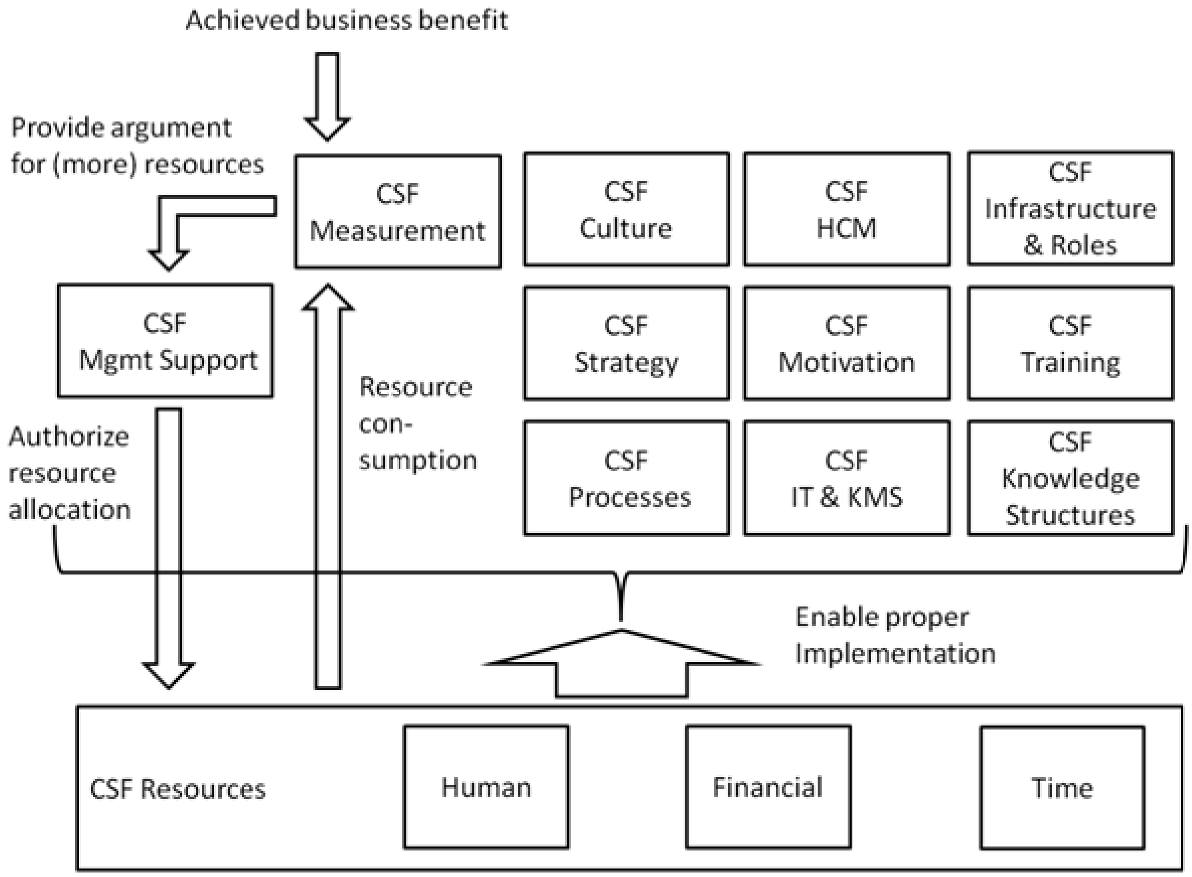


Figure 4-8 ***Relationship between CSF Resources and the other CSF***
Source: Own illustration

Cross-functional activities, such as KM initiatives, are often funded through an organization's overhead cost collection. If this financial pool is small or if other pressing initiatives rival for the allocation of money from it, it is of the essence to focus the KM initiative on smaller scale KM projects that will target the most urgent knowledge-related problem areas and needs of the organization. These areas should promise quick-wins. When KM is properly introduced in these areas, the company can expect the biggest impact for the organization and can help building up future trust in the value of KM projects and hence increase the potential to allocate money from overhead (Davenport et al. 1998).

Smaller scale approaches are especially important to *Small and Medium Sized Enterprises* (SMEs) (Wong, Aspinwall 2004, 2005). These companies have to consider KM initiatives with special care, since “SMEs are not like ‘little large businesses’” (Wong 2005, p. 266). In general, their financial funds are substantially more limited and they have a smaller pool of human resources to fill up KM-related roles and positions. Additionally, these companies often cannot rely on a strong financial overhead generated by other departments. Consequently, the relative cost of KM in contrast to turnover and profit is much higher and SMEs cannot afford investing in KM as much as larger companies can. They need to consider the scope of KM programs in accordance to their scarce resources and limit it to an extent that is feasible within their situations.

Summarizing the discussion, sufficient provision of resources is essential to the success of an initiative. The resources are used to support the achievement of satisfactory results in the other CSF. Resources are always scarce which is why *top management support* is necessary for their allocation to the KM initiative. If no dependable figures can be generated that indicate the extent of (financial) benefit for the company, top management support is of utmost importance to even start the initiative. If a measurement system is in place that can verify the benefit of the initiative, it can help to provide arguments for management to (further) support the initiative. If resources are very limited, the KM initiative should focus on quick-win areas that demonstrate benefit as fast as possible to have arguments for extending the initiative’s scope. This is especially true for SMEs where KM initiatives have a higher relative cost than in larger companies.

4.2.15 Summary of CSF for KM Initiatives

The previous sections described and analyzed the available literature to extract CSF generally relevant for KM initiatives’ success. Subsequently, the interpretation of each CSF was explained in detail by further analyzing the literature for each CSF. Twelve CSFs were identified in the process that need to be taken into account for a KM initiative to be successful. Along with the description of each success factor’s interpretation, also the interdependencies between the CSFs was addressed. This highlights the complex nature of KM initiatives in which all of the twelve CSFs need to be in concert for overall success. While the dependencies between the CSF were highlighted where appropriate, the exact relationships between the factors remain unclear. Additionally, there was no indication to be found on the relative importance among the CSFs let alone empirically deduced quantitative results as to their exact influence. So far, the state of the art in CSF research for KM initiatives could only produce lists of factors that are relevant. While this is very essential knowledge by itself, a structuring of the CSF still promises benefits especially for practitioners and also is a first step into demonstrating at least qualitatively the relationship between the CSF. Other authors have addressed this necessity already. Helm et al. (2007) propose a model in which the CSFs are arranged in the four dimensions *personnel*, *culture*, *structure*, and *KM activities*. They also formulate relationships between those dimensions. However, as they state themselves, some of their dimensions, for example *culture*, appear to be endogen dimensions, i.e. they can only indirectly be influenced by changes in the other dimensions. Additionally, Helm et al. (2007) consider IT as a sub-dimension of their dimension *structure* that also comprises organizational structure. IT in general and KMS in particular form the backbone of most KM initiatives. It is, therefore, imperative for organizations to be aware of the capabilities of these systems and their technological advancement (Moffett et al. 2002). For new KM initiatives, this makes decisions on implementation options less straightforward.

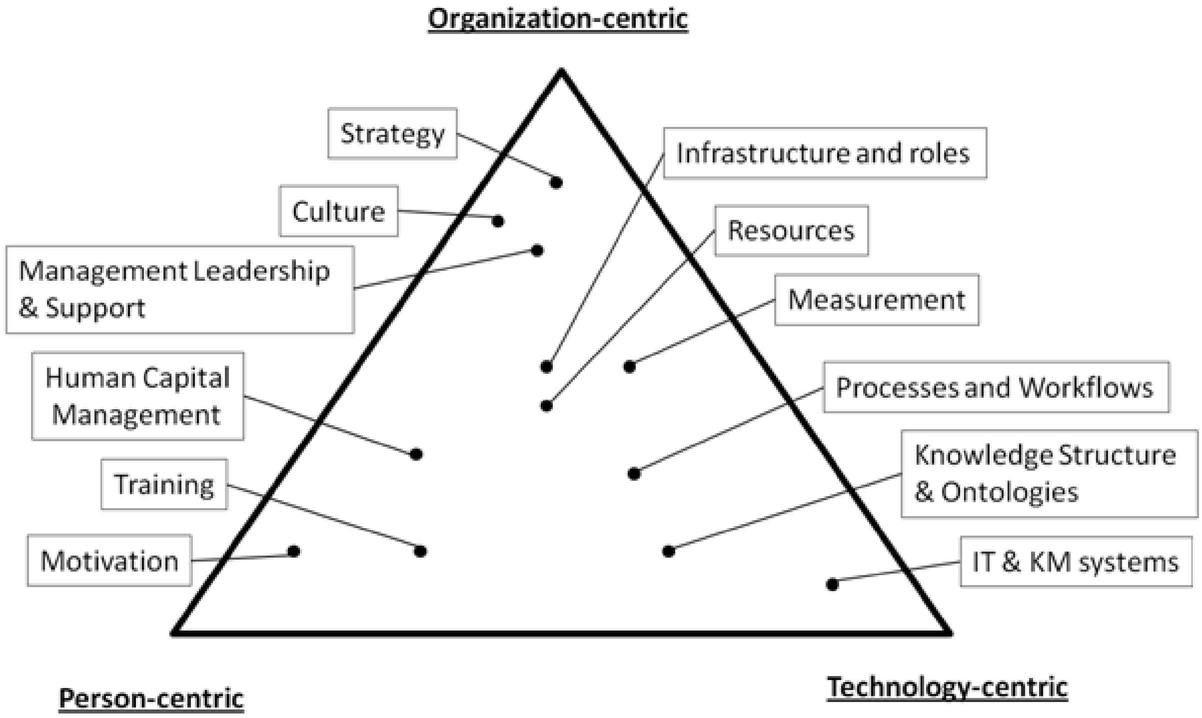


Figure 4-9 ***The twelve CSFs and their relation to the personal, organizational and technological dimensions***
Source: adapted from (Schmidl et al. 2011a)

In this thesis, a different approach to structure the identified CSFs is taken. In KM, the three focus areas *person*, *organization*, and *technology* have often been discussed as cornerstones of KM initiatives (Bullinger et al. 1997; Bullinger, Wildemann 2002; Lehner 2009; Krcmar 2010). The wide adoption throughout the community renders those perfect candidates for structuring CSFs. Furthermore, common agreement starts to form that KM initiatives are combined social, organizational and technological endeavors (Hasler Roumois 2007). Additionally, in a company, these dimensions can be mapped to specific areas. While technology-centric factors are likely to be implemented by an organization's IT department, the organization-centric tasks are rather to be supported by (senior) management. In contrast, person-centric success factors strongly depend on the operational level of an organization. Following this argumentation, the CSFs identified in this chapter are structured according to these three dimensions as depicted in Figure 4-9. However, no CSF is independent of the others and neither can one CSF only be associated to one of the three dimensions. Therefore the success factors are assigned to those parts of the triangle that are most influential for their proper implementation. Knowing which part(s) of an organization is most influential for a certain success factor helps practitioners to understand where to invest the most effort. This way, the proposed structure of CSFs offers practitioners a direct benefit. Knowing which part of an organization is most influential for a certain CSF helps them to understand what to focus on and also gives insights for the design of new KM initiatives.

4.3 Success Factors for IT-based KM

IT or more specifically KMS are the backbone of many KM initiatives, especially when limits of space and time need to be bridged. This is why organizations should be aware of the capabilities of these systems and their technological advancement (Moffett et al. 2002) when they want to introduce a KM initiative. This thesis' KM approach will also rely on IT to bridge the limits of space and time. This is why in this section, the elicitation of key features and properties that have been reported to make KMS implementations successful is addressed. Publicized case studies about KM project will form the basis for the analysis. The focus will be on those KM case studies that contain sufficiently elaborate description of the implementation of KMS. Further, the focus will be on *Knowledge Repositories*, *Expert Directories*, and *Communities of Practice*, as these have been reported to be the KMS that are most often introduced in companies, which promises to deliver more reliable results (Alavi, Leidner 2001; Riempp 2004; Laudon, Laudon 2006; Krcmar 2010). In the following, first the survey sources are elaborated on, before each identified CSF for IT-based KM is detailed.

4.3.1 Approach for Determining KMS SuccessFfactors

A wide range of KM case studies has been publicized. In order to scan and examine a large number of case study sources in an efficient manner, this review considers primarily case studies that have been published as collections in monographs. The Gateway Bayern - Bibliotheksverbund Bayern (BVB[4]) electronic library was used as main research platform for identifying books and other sources containing case studies. Additional sources were determined by following references included in the case study articles. Table 4-6 depicts the main sources examined during the case study analysis ordered by year of publication.

[4] The BVB is a regional cooperation project and connects over 100 public and university libraries in Bavaria, Germany

Citation	Title
Gottschalk (2000)	Knowledge Management Systems: A Comparison of Law Firms and Consulting Firms
Blessing (2001)	Content Management für das Business Engineering
Eppler (2001)	Fallstudien zum WM: Lösungen aus der Praxis aufbereitet für die Aus- und Weiterbildung
Nonaka (2001)	Knowledge emergence: social, technical, and evolutionary dimensions of knowledge creation
Abecker et al. (2002)	Geschäftsprozessorientiertes Wissensmanagement effektive Wissens-nutzung bei der Planung und Umsetzung von Geschäftsprozessen
Barnes (2002)	KM systems theory and practice
Bellmann, Krcmar, Sommerlatte (2002)	Praxishandbuch Wissensmanagement Strategien - Methoden - Fallbei-spiele
Bullinger, Wildemann (2002)	Wissensmanagement Wissen als strategische Ressource im Unterneh-men
Sukowski (2002)	Knowledge management case studies project experiences, implementa-tion insights, key questions
Davenport, Probst (2002)	KM case book Siemens best practices
Mertins (2003)	Knowledge management concepts and best practices
Rollett (2003)	Knowledge management processes and technologies
Koenig and Srikantaiah (2004)	Knowledge Management lessons learned what works and what doesn't
Dixon (2005)	Common knowledge how companies thrive by sharing what they know
North (2005)	Wissensorientierte Unternehmensführung Wertschöpfung durch Wis-sen
Kazi, Wolf (2006)	Real-Life Knowledge Management: Lessons from the Field
Probst, Raub, Romhardt (2006)	Wissen managen wie Unternehmen ihre wertvollste Ressource optimal nutzen
Maier (2007)	Knowledge management systems information and communication technologies for knowledge management
Völker, Sauer, Si-mon (2007)	Wissensmanagement im Innovationsprozess
Nonaka, Toyama, Hirata (2008)	Managing flow a process theory of the knowledge-based firm
Laudon, Laudon (2006, 2007, 2009)	Management information systems managing the digital firm
Lehner (2006, 2008, 2009)	Wissensmanagement Grundlagen, Methoden und technische Unter-stützung

Table 4-6 ***Analyzed case studies to determine KMS success factors***
Source: Own illustration

The scope of the analysis was limited to include the last ten years, which roughly marks the rise of the *third generation of KM* (Schütt 2003).

The analyzed sources cover a large variety of businesses and industries which for instance include electronics, mechanical engineering, software development firms, professional services like consulting and auditing, automobile manufacturers, aircraft manufacturers, financial services, banks, oil and gas as well as specialty chemicals corporations, office equipment, pharmaceuticals, health care firms, and defense companies. The number of the employees in the examined companies ranges between 50 people to global corporations that currently employ up to over 180.000 employees.

Some of the more publicly known enterprises studied during the research included: Hewlett Packard, Xerox, Ford, Nokia, Accenture, Roland Berger, PricewaterhouseCoopers, Honda, Volvo, Siemens, Dow Chemicals, AT&T, General Electric, SD&M, Aventis, Infineon, Ernst&Young, IBM and Texas Instruments.

Despite the big range of examined companies and industries all the KM case studies fall into relatively few categories and can be most easily structured around their main goals (Sukowski 2002, p. 16). The review revealed that organizations largely perform their KM projects in pursuit of four general types of objectives that include:

- KM projects that focus on implementing and introducing new KMS.
- Initiatives that solely focus on facilitating knowledge transfer through direct contact or for instance by deploying videoconferencing.
- Projects that target the exploitation and measuring of the value of knowledge in order to manage and exploit it as an asset.
- KM culture campaigns that are geared towards improving the overall environment, awareness, and social conditions for managing knowledge.

Although some KM initiatives had a broader scope and attempted to attain several of the goals simultaneously, most of the case studies were concerned with achieving one goal only. As this chapter’s goal is to determine success factors for IT and KMS the further analysis of the literature review focused on analyzing case studies that had the implementation of KMS as their primary goal. Only case studies that were sufficiently elaborate in their description of the KMS’s features and the implementation procedure were included. To further focus the analysis the analysis was restricted to implementations of *Knowledge Repositories*, *Expert Directories*, and *Communities of Practice*, as these have been reported to be among the first systems to be introduced in KM initiatives (Alavi, Leidner 2001; Riempp 2004; Laudon, Laudon 2006; Krcmar 2010). While not true in all cases, focusing on these systems promises to give a more detailed insight in companies that start with KM initiatives rather than those that extend their KM initiatives with further projects.

In the following, the eleven factors that were extracted from the case studies are explained. Since KMS are special kinds of IT systems, considerations that hold for other IT systems will also arise as success factors during the analysis. While general CSF for KM were already analyzed in sub-chapter 4.2 and the focus now lies on the technological support, some overlap with the general success factors are natural as none of the CSFs of a KM initiative is independent of the others (see 4.2.15).

4.3.2 Critical Success Factor 1 – Selection of Software Platform

Organizations that implement KMS typically impart their use into their long-term strategy, i.e. the KMS are intended to be used over an extended period of time, which is typical for organization-wide IT systems. It is, therefore, important for organizations to choose IS and technologies as platforms that will guarantee the longevity, stability, and availability of the organization's knowledge in a flexible manner. A proper selection of software is especially important as replacements of organization-wide systems are typically complex and resource-intensive (Stahlknecht, Hasenkamp 2005) also known as lock-in effect. In a first instance, organizations have the choice of "make-or-buy", i.e. to either create a *specifically tailored* solution for the own organization or to procure *standard software* (Schwarzer, Krcmar 2004, p. 224). Standard software has some advantages but also disadvantages in comparison to the development of individual software that an organization needs to carefully weigh against each other. The most important aspects are summarized in Table 4-7.

Advantages	Disadvantages
Cost savings through economies of scale at the provider and larger group of clients	Incomplete fulfillment of company-specific requirements
Eliminates development times; solution available immediately or very quickly	Limited integration into the company's application portfolio, e.g. due to interface challenges
Reduction of introduction or change times in comparison to individual software that is often implemented module by module	Due to focus on general usability, potential of suboptimal properties in company-specific context
High quality of programs due to high specialization of providers and potential competition between providers	Only the effort for one part of the application development cycle can be saved, as there is still effort for adaptation and introduction
Application maintenance and further development is guaranteed by the provider	
Application development is independent of the company's available IT personnel	

Table 4-7 ***Advantages and disadvantages of standard software in comparison to individual software***
Source: Adapted from (Krcmar 2010, p. 170)

The results of analyzing the case studies indicate that the selection and later adaptation of standard systems promises greater benefit than disadvantage. The standard system however has to fulfill the minimum organizational requirements. One important advantage is their support for a substantial number of interfaces to other (source) systems as well as migration and extension possibilities for connecting additional external solutions or adapting current ones down the road (Brügge, Dutoit 2004). Along with the reduction in cost for own implementation, these possibilities can save organizations a lot of time and money when they want to extend functionality in the future or have to merge or exchange knowledge with other organizations. Standard solutions are also more likely to have a bigger base of knowledgeable experts and if developed by an established and secure software/hardware producer, the risks of having supporting issues in the future are small. Finally, Krcmar (2010, p. 170) argues that standard software also tends to be cheaper than development of own solutions – a very important argument for IT-supported KM initiatives. The value of KM initiatives is hard to determine (see 4.2.7) but costs are obvious. If costs are high while the benefit is not as obvious, the risk of the KM initiative losing support is high, which is why costs should be kept small especially in the beginning of a KM initiative.

4.3.3 Critical Success Factor 2 – Choosing the Right System

Selecting an appropriate IT solution is a typical task in *information management* and there is a variety of approaches. Krcmar (2010, pp. 171–178) gives a good overview of them. In addition to the general considerations, in the following those aspects for choosing the right KMS are addressed that are either specific for KMS or especially important in their context.

An organization typically favors one of the two stereotypic *KM strategies*: codify knowledge for reuse or support the transfer of knowledge from individual to individual (see 4.2.3). Consequently, KMS should be in line with this strategic choice. Furthermore, they should first address the most glaring needs of the organization as a whole or of the main business processes that they integrate into. However, the strategic needs of the enterprise are not the only major focal point during the decision process that determines which KMS will create most benefit for the organization. The *organizational culture* (see 4.2.5) should have a significant influence on the selection of a solution as well. For example, if a company's employees are used to finding solutions very quickly and this is what they believe their work needs to be, a KMS that relies on extensive, manual documentation of the employees' solutions extends the solution finding time and, hence, is against the current culture in the company. Another important aspect is to create and implement KMS that yield benefit to the employees using the KMS, i.e. the employee must be *motivated* to use the system (see 4.2.11). For example, systems should not undermine the status and expertise of employees and have to be in accordance to their skill-level to avoid frustrating user experience. They should rather provide obvious benefit for the employee like easier processing of operational tasks. Neglecting the company culture and motivational aspects can be a costly mistake. Projects that solely focus on building systems out of strategic need, without considering cultural characteristics, are likely to fail. The assumption that "if we build it they will come" has proven to be a dangerous one (O'Dell, Grayson 1998). In the case studies, it turned out that some organizations still introduced KMS that were not fully in line with company culture but rather with the (future) KM strategy. In these cases, employees were less willing to adjust their habits if they feel that the adjustments are proclaimed as necessary just because management has spent a lot of money and now "forces" a new, perhaps in their eyes unnecessary, technological solution upon them. Marketing the KMS in conjunction with training measures is therefore especially important (see 4.2.12). Furthermore, for KM initiatives to be successful in this context, it is important to either address and possibly adjust cultural aspects before a system is introduced or to choose a system that might not be the ideal strategic solution but will fit in with existing corporate culture. Adjusting the culture, however, is a task that is hard to perform and that takes considerable time. Additionally, the argument, that a culture might be wrong and needs to be changed for a KMS to be effective is disputed in literature (Schütt 2003, p. 453).

4.3.4 Critical Success Factor 3 – Integration with Existing Business Processes

Ideally, employees should have seamless access to the necessary knowledge when performing their daily operational business tasks. If, for example, a KMS offers the ability to automate parts of the knowledge retrieval process such as simplifying search processes it will fulfill this property. More generally, if some sort of automation can be incorporated into existing tasks and/or currently used technologies of the organization, it supports the perception that KM is just a *by-product* of the daily work and not an *extra burden*. KM efforts will hence be (partially) masked and the productivity of employees can be increased by saving time (see 4.2.9). The integration between KM and operational business processes can even form the focal point for KM initiatives (Heisig 2005). Tight integration of operational business processes with KMSs' functionality, hence, proves to be an important aspect of KMS implementation.

4.3.5 Critical Success Factor 4 – Functionalities for Content Management

Organizations face the challenge to make sure that the knowledge contained in their KMS and provided to their employees is valid and adheres to certain quality standards. Some means of quality assurance is therefore necessary. However, the measure depends on the kind of system used. In Knowledge Repositories, often dedicated personnel enacts a publishing review workflow and might also perform audits of older items in the knowledge repository. In CoPs, on the other hand, the community itself governs what is valuable. Especially since the widespread adoption of tools form the Web2.0 into the enterprise world such as wikis, blogs and diverse forms of online communities, user ratings serve the purpose of indicating the relevance and quality of knowledge. Still other approaches take e.g. the usage and usage context of a knowledge resource as a proxy for its relevance. For example, many search engines operate like that when taking the contextual information contained in the link structure as proxy for the relevance of a document.

4.3.6 Critical Success Factor 5 – Support of Individual Evaluation of Knowledge Resources

Over time, the amount of knowledge that is accessible to a company's employees can be substantial and can start to cause a new problem: Judging which knowledge is suitable, i.e. relevant and applicable to the current situation becomes increasingly complex. If the effort for finding suitable content starts to outweigh the benefit of the knowledge resource, the usage of the KMS starts to dwindle and the benefit of the KMS will likewise. KMS should, therefore, offer some means to allow its users to evaluate the fit of a knowledge resource to the current situation of the user. In addition, they should ideally also rank the most suitable knowledge pieces to reduce the *cognitive load* for the user. Those ranking functionalities allow employees to filter large amounts of knowledge more effectively during their searches and help organizations to gain insights onto the quality of the knowledge resources available in the organizational knowledge base (see previous section). In addition, it is also possible to use rating functionalities for tracking the quality of contributions that individual employees have made to the organization. Previously defined quality characteristics and results can subsequently be used as a basis for the performance evaluation of employees and allow identification of the "best" knowledge contributors within the organization – a *motivational* aspect that can positively influence knowledge sharing behavior (see 4.2.11). An additional mechanism that some KMS offer is *content push*. When employees actively search for content, they *pull* content when they need it. However, a co-worker might have shared knowledge that is important for another employee but not at the very moment. Some systems allow to *push* this new content to the employees that will likely value the additional knowledge – a form of precomputed, individual evaluation of knowledge resources.

4.3.7 Critical Success Factor 6 – Support for Navigation in Knowledge Structures

When retrieving pieces of knowledge, there are two typical approaches. Some users rely on search applications, others are more used to navigating a logical structure to find what they search. An analogy would be a person trying to find a file on his computer: Some users rely on desktop search engines or their operation system's search facilities to retrieve files, while others rather browse the folder structure of their file system to find the desired file. The first case has been addressed in the previous section. The second case can be interpreted more broadly. Pieces of knowledge can be included in a knowledge structure such as an *ontology* (see 4.2.8) and KMS should support the creation and maintenance as well as the navigation within these knowledge structures. The terms that are used for classifying a piece of knowledge can either be selected from a company-defined set that is centrally maintained or, as it is the case for applications like wikis or blogs, can be assigned user-defined tags.

4.3.8 Critical Success Factor 7 – Roles and Rights Management

An organization's unique knowledge may create a competitive advantage on the market and, therefore, is precious. Similarly, some knowledge in an organization, should be available only to specific employees, not all. For example, discussions among senior managers in a company's forum about the long-term strategy of the company should not be visible to anyone outside this group. For the individual employee it might also be imperative to know who can access the knowledge that he shares. Some pieces of knowledge might not be shared if the individual employee cannot trust that the piece of knowledge cannot fall into the wrong hands. Therefore, a KMS needs to be able to determine who is responsible for certain knowledge resources and which members of the enterprise are authorized to view, create, and modify particular content, i.e. it needs to be able to reflect a company's *governance structure* for KM (see 4.2.6). The integration of existing identity management solutions within the company is a straightforward way, but it can also be possible that the KMS supports its own roles and rights management system by distinguishing individuals and groups.

4.3.9 Critical Success Factor 8 – Appropriate User Interface

The use of any software depends among other factors on the *ease of use* (Davis 1986, 1989) or its *expected effort* (Venkatesh et al. 2003), respectively. Therefore, companies should select KMSs that are in accordance with the employees' "computer literacy", i.e. their skill level concerning the use of IS. Choosing systems that rely on standard, intuitive, and user-friendly interfaces that are similar to ones that the employees know from their everyday tasks, increases the chances of the KMS being accepted and applied correctly and successfully in a more frequent manner. Depending on the type of KMS, there are some features such as context-dependent frequently asked questions (FAQs) documents, online user manuals, dynamic help, provision of support contact information and "what you see is what you get" (WYSIWYG) editors to increase the ease of use for the employees (Maier 2007; Lehner 2009). Providing some type of personalization is another possibility because it allows employee to reflect in the interface their way of thinking, organizing, and navigating through the existing organizational knowledge.

4.3.10 Critical Success Factor 9 – Availability and Accessibility

In correspondence to the logistic principle for information (Augustin 1990), the goal for KM should be to offer the *right* knowledge at the *right* time in the *right* quantity at the *right* place at the *right* quality. The determination of what is the right knowledge, the right quantity and the right quality of knowledge has been discussed above (see 4.3.5, 4.3.6 and 4.3.7). The right time is whenever the employees needs a piece of knowledge for performing a task. This is why the availability of a KMS is essential – the logistic principle loses its value if the system is not available when needed. This is true for other kinds of IT systems as well, but KMS are often support systems. In contrast to operational systems, the company's employees have a larger degree of freedom to not use the KMS and hence availability is more important to have the employees *trust* in the KMS. This is especially important during the introduction of the KMS where it has to prove its added benefit to the employees to be accepted.

The second aspect of the logistic principle that has not been addressed yet is "at the right place". In addition to being available and operational, KMSs need to be available at the right place of knowledge application and usage. For example, if a KMS is accessible only from within the companies intranet and an employee is working outside of it without being able to access it from the outside, he will not be able use the KMS although it is up and running. It is not available at his place of knowledge application, which erodes the benefit it could provide. Organizations, therefore, have to carefully consider their requirements when analyzing possible solutions and for example take into account features like remote access, support of mobile devices and handhelds, and offline usage via checking in and out while still keeping an eye on other corporate necessities like IT security measures.

4.3.11 Critical Success Factor 10 – Maintenance Efforts

During their life time, 70 % of the overall costs of IS are spend on maintenance (Erlikh 2002). In a survey, CIOs indicated that growth (of the IT system and the content they manage) is the primary cost driver for enhancing existing IT systems and further integrating them in the company infrastructure after initial launch, i.e. in the maintenance phase (Accenture 2005, p. 6). Additionally, they note that the number of errors is the primary cost driver for typical post-implementation tasks such as fixing problems and testing changes. This is not different for KMS. However, while operational systems are often virtually irreplaceable, KMS could be abandoned if the costs increase too much, while, as argued already, the benefit is not as easily measurable (see 4.2.7). Though the complexity, scalability, performance, reliability and dependability as well as the resulting maintenance efforts of IT systems should always also have an influence on the selection of an appropriate solution (Krcmar 2010, pp. 173-178}, the relevance of these aspects for KMS is arguably larger.

4.3.12 Critical Success Factor 11 – Implementation Approach

IT systems can be introduced in different ways. Daniel distinguishes three archetypes of implementation approaches that differ in the following six implementation aspects (Daniel 2001):

- Degree of end user participation
- Scope of IT solution
- Degree of perfection
- Scope within company
- Contextual overlap
- Time of implementation

The first archetype, roughly translated *big bang* promises short implementation times but is the most risky approach which, in order to control, leads to extensive planning before the implementation. Further, much experience in the implementation of similar applications is necessary. In addition, the acceptance of KMS cannot be anticipated very accurately. Finally, experience in implementing KMS is hard to obtain as the contextual factors, such as the corporate culture are on the one hand very influential for the success and on the other hand different for every project. Therefore, it is not advisable to perform KMS implementations in a big bang approach.

The second archetype that Daniel (2001) identified is roughly translated the *emergency project* approach. It focuses on quick wins due to high time and cost pressure, all other aspects come second. Quick wins are also a goal for first implementations of KMS to quickly prove its value and get further support and resources. However, in emergency projects, solutions are implemented as soon as possible and not once it is sufficiently mature. This lack of maturity increases the risk to erode *trust* in the KMS, which is especially problematic for KMS as their use relies more on the employees' trust and willingness to adopt. Furthermore, this kind of implementation project is implemented in a non-participatory manner to ensure quick results. Yet, integrating the users into design and implementation decisions is often a crucial aspect of KMS.

Daniel's third archetype, the *slow optimization* is suitable if quality and acceptance is most important, but may only work if there is no high time pressure (Daniel 2001, p. 177). Therefore, if no contextual influence like time pressure, force the company's decision for an implementation strategy, the third type of implementation seems most suitable for KMS.

4.3.13 Summary of Success Factors for KMS

In the previous sections, the analysis of the literature to extract CSF for KMS was described. KMS in turn contribute to the success of the KM initiatives in which they are used. Subsequently, the interpretation of each success factor was detailed. Eleven success factors could be identified that should be kept in mind when a KMS' support for a KM initiative is to be successful. This provided some insight in the technological dimension of KM initiative's success. Because no dimension can be considered independent of the others (organization and person; see 4.2.15), along with the description of each success factor's interpretation also the interdependencies to the general CSFs of KM initiatives were addressed. This way, the complex interactions of technology as enabler of KM initiatives and the remaining factors for success were highlighted. Figure 4-10 shows an overview of the KMS success factors. The *choice of system type* is done as a first step, taking into account the organization's context (*culture*, *strategy*, way of *motivating*). It gives indications which *software platforms* may be usable. The eight upper success features indicate the requirements that need to be fulfilled by the selected software platform. The left hand success factor i*mplementation approach* spans across the others. While the remaining ten list features of the system, the implementation approach deals with the introduction process itself. The right hand site shows which of the general CSFs (see 4.2.15) interact most with the respective KMS success factors.

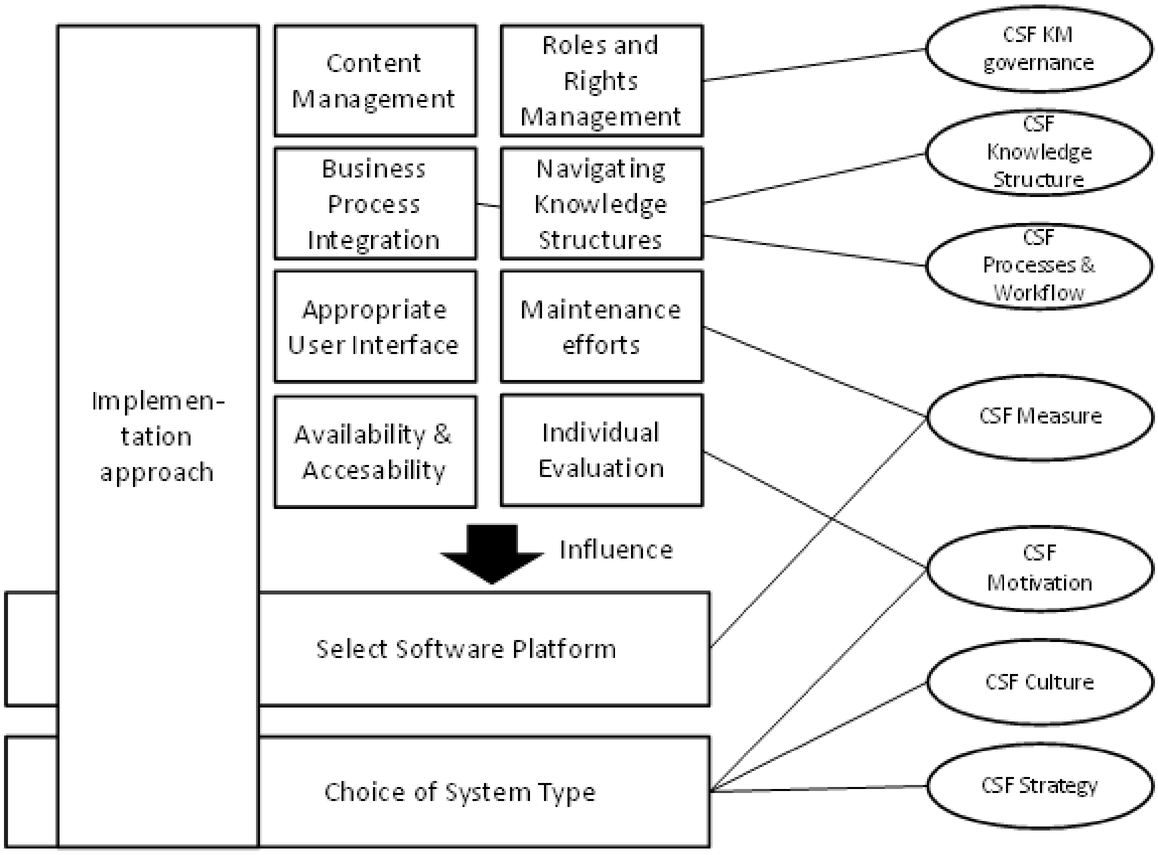


Figure 4-10 ***Success Factors for KMS, their relationship among each other and the connection to general KM initiatives' CSFs***
Source: Own Illustration

4.4 Summary and Discussion

In this chapter, the CSFs that predominantly determine the success of a KM initiative were analyzed to understand which aspects have to receive special emphasis, when later developing this thesis' KM approach. A common set of twelve CSFs that have to be managed well in any KM initiative, were determined by analyzing the current literature on KM success and failure. For each of the factors a more elaborate literature search was performed to find sources that detail the interpretation of the factor. This was necessary as the factors were often only shortly scratched in the main sources that were used to determine the success factors. To bring the factors in a more general relation, the determined factors were arranged in a structure relying on an often used three-dimensional categorization in KM, which uses the dimensions *person, organization* and *technology*. Figure 4-11 depicts the final categorization.

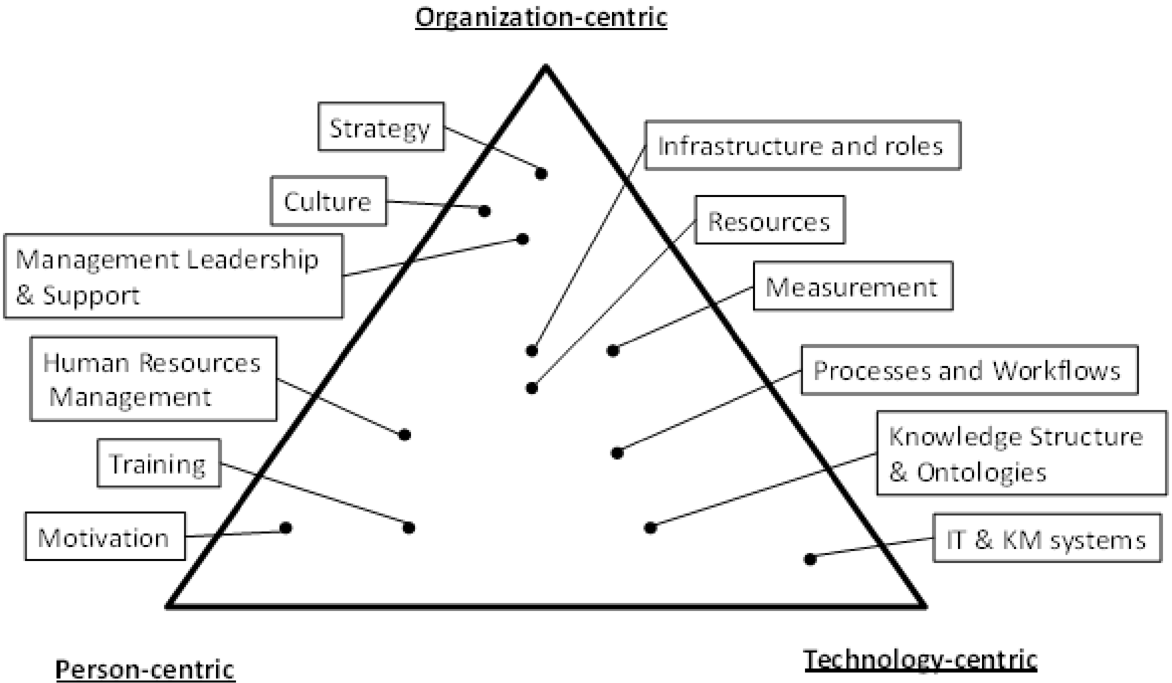


Figure 4-11 ***The twelve CSFs and their relation to the personal, organizational and technological dimensions***
Source: (Schmidl et al. 2011a) and chapter 4.2.15

Additionally, as this thesis' KM approach will rely on IT support, a more focused literature review was conducted that determined the success factors specifically for KMS. Here, a set of eleven factors that influence the success of a KMS could be retrieved. The factors along with their relation to the general KM initiative's CSFs are depicted in Figure 4-10.

Knowing the deficits of contemporary KM approaches that support the socialization process (see chapter 3) and now having an understanding of which aspects need to be treated with special care when implementing a KM initiative, provides the necessary conditions to develop this thesis' KM approach in the following chapter.

5 Concept for Fostering Knowledge Exchange Based on Task Overlap

"Ähnlichkeit ist die Identität der Qualitäten"
Freely translated: "Similarity is the identity of qualities"
Immanuel Kant (1724 - 1804), German philosopher

Ackermann and Halverson argue that *Knowledge Management* is not the most suitable term for the pressing challenges of dealing with knowledge anymore. They think that " […] we move from the metaphor of *knowledge management* to a new metaphor, *expertise sharing*, which promotes focusing on the inherently collaborative and social nature of the problem." (Ackerman, Halverson 2004, p. 1). Independent of the KM strategy that a company follows (see 4.2.3), the social setting of a company is the basis of any attempt to improve the flow of knowledge within a company.

In this chapter, this thesis' KM approach will be developed. It builds upon *task similarity* as its driving force. It will combine features of the approaches that were investigated in chapter 3, while trying to circumvent their disadvantages. Additionally, care will be taken to adequately address the influential factors that were elicited in chapter 4. Yet, the design will not start on the playground from a research perspective. Therefore, this chapter also investigates possibly suitable models from theory that will inform the design of the KM approach.

As a first step, design considerations that influence the further development will be discussed. In the second part of this chapter, theoretical insights will be investigated to understand implications for this thesis' design. Then, an overview of this thesis' approach will be provided before each layer of the resulting model will be explained in elaborate detail.

5.1 Preliminary Design Considerations

The analysis of currently employed approaches for fostering the direct exchange of knowledge between individuals in chapter 3 has identified specific deficits for each. Looking at the results more closely one can see that, while there are some deficits that apply to all approaches, there are some challenges that only apply to YPS and ERS respectively, and there are other challenges that only apply to knowledge networks. For example, the challenge to increase the number of individuals that benefit from interactions is inherent in systems that are designed to establish one-on-one connections like YPS and ERS, but is not a problem for knowledge networks that are established and nurtured to bring together a larger group of individuals for mutual learning. On the other hand, the challenge of establishing new or finding suitable existing communities for personal engagement obviously applies for knowledge networks, but the very idea of YPS and ERS is to facilitate the process of finding suitable interaction partners and hence the challenge is less expressed there. It therefore seems promising to exploit the respective other's approach's benefits in a *combined approach* for fostering knowledge exchange. As Figure 5-1 shows, this is exactly what this thesis aims at. However, there needs to be a suitable "glue" to combine the two approaches – knowledge networks on the one side and ERS on the other. There needs to be a connecting link that, on the one hand, can serve as basis for *deriving expertise* – the goal and inner working of ERS – and, on the other hand, can serve as basis for *bringing like-minded individuals together* – the fundament of knowledge networks. This thesis' approach will use the *similarity of tasks* as its glue. The (repeated) execution of tasks is a *hint for expertise* for the work that is executed in this task (Greer et al. 1998a, p. 166; Terveen, Hill 2002, p. 10; Seid, Kobsa 2003, p. 13). At the same time, the executed tasks give an indication of the kind of occupation of an individual and, hence, allow to find others that are *similar* with respect to their *task and occupation*, which provides common grounds in terms of *shared values*, *norms*, *perspectives*, and *language* (van Maanen, Barley 1984, p. 287; Wenger et al. 2002, pp. 29–32; Teigland 2003, p. 35) and hence facilitates knowledge transfer (Bromme et al. 2004).

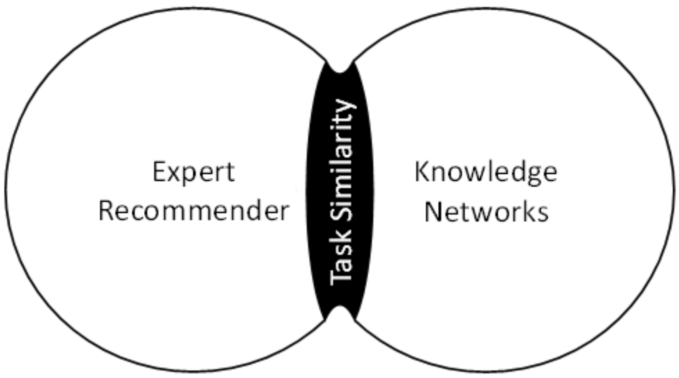


Figure 5-1 ***The context of this thesis' KM approach***
Source: Own illustration

The advantage of knowledge networks as well as mechanisms for relating single individuals as approaches to improve knowledge flows in organizations has been emphasized before. For example, Ackermann and Halverson state that an organization's expertise sharing can be improved by both – by routing queries to individuals with sufficient expertise or by creating places for people with suitable expertise to meet and others that can join and learn from them (Ackerman, Halverson 2004, p. 13). While they seem to indicate that these two approaches are disjoint, this thesis' argues that a combination of both into one approach is advantageous and possible.

The possibility to integrate functionalities of knowledge networks and recommendation systems has also been taken into account before. For example, Terveen et al. (2002) note that recommender systems that deduce an individual's preferences by observation of their natural behavior can create opportunities for community building. While just stating as a suggestion and not giving further leads to a possible solution, they indicate that people who share preferences could be put in touch (Terveen, Hill 2002, p. 11). In contrast to their suggestion, this thesis does not use preference measures but *executed tasks* as its basis for relating individuals and shows an implementable approach to reach this goal.

In pursuit of this goal, some, on occasion pragmatic, conditions that this thesis' approach needs to fulfill, need to be considered. It is necessary to carve out on *which tasks* this thesis' approach will focus, and how to deal with *anonymity* that arises in distributed settings to a more or less pronounced level. In addition, the particularities of the approaches that are incorporated need to be considered and experiences from their implementations should enrich the design of this thesis' approach. Finally, pragmatic considerations like *embedding* the solution into operational systems and the handling of possible *information overload* needs to be considered. Each of these topics will in turn be addressed in the following.

5.1.1 Focus on IT-Related Tasks

Nowadays, large portions of an organization's operations are supported by IS. In a survey and log file analysis of knowledge workers in a governmental unit, Makolm et al. (2007) found that roughly ***three quarters*** of all activities were supported by IS, while activities like face-to-face or phone discussions, meetings and mental work was not. IS use can be evidenced along the whole value chain. Supplier Relationship Management systems help maintaining beneficial contacts with suppliers, customer relationship management systems support the other end of the value chain keeping track of the contacts and interactions with customers. In between there are systems that support financial accounting and that keep track of projects. Other more general systems, like workflow management systems automate process executions to a sometimes very high degree and ERP systems form the backbone of the multitude of operational systems in an organization. Especially knowledge-intensive tasks are performed in the digital world to a large degree. Finding information through various sources, (re-)combining it and placing the results "back into the process", i.e. into IS or digital storages is a typical process for knowledge workers. In general, this process is not at all or only to a small degree a structured process and the knowledge worker decides which single steps to follow to achieve his goal in this process.

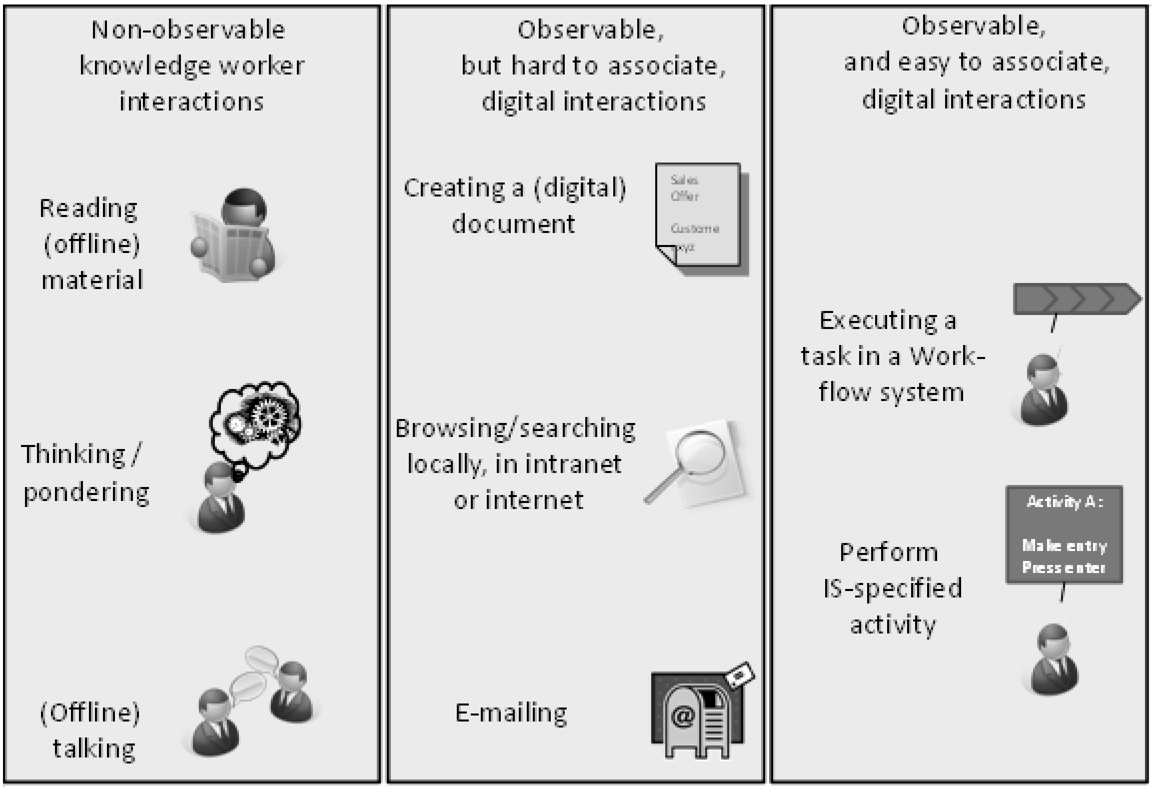


Figure 5-2 ***Classification and examples of typical knowledge worker activities***
Source: Own illustration

A knowledge worker may perform diverse activities during his workday as illustrated in Figure 5-2. For example, information is gathered through *offline reading*, new information is created through an individual's *thinking*, and mentally recombining what the individual knows or by *talking* to other coworkers and peers. While there are exceptions e.g. (Wu et al. 2008), those interactions are normally not observable by IT systems and may often not be observed at all, due to technical, social and legal limitations. However, other ways of gathering and storing information can be observed by IT systems, such as *searching* in the company's intranet, in the internet or on the local file system. Likewise, the *creation of documents* such as text files or presentations can be observed which is also possible for *computer-mediated communication* such as e-mail messaging or instant messaging. However, it is a difficult task to understand why the knowledge worker does these activities, i.e. which results are pursued while performing the activities. For example, when a knowledge worker writes an e-mail it is not directly obvious what it is about. Content analysis using text mining procedures might shed some light on its content but still the reason and exact content remains only vaguely interpretable. However, the results of the knowledge worker's mental processing is eventually integrated into the company's business processes; it is entered into IS using this system's functionality and hence the IS support for tasks. On the other hand, if a more formalized and IT-supported process is enacted, the knowledge worker's efforts result in the completion of a task in a workflow system. For example, a sales offer might be stored in the customer relationship management system after executing a "send sales offer to customer" task in the ERP system. Likewise, the marketing material might be stored somewhere on the intranet's folders, after executing a "plan marketing campaign" task in the customer relationship management system. These activities *are* observable. In addition, since they are related to distinguished functionalities in the system, and, if one exists, also to a business process model, the execution of the activity can be *associated with its goal*. Additionally, one can reason that executing these tasks allows inferring the executing individual's ability to gather the necessary information using digital or non-digital interactions and hence one can find evidence for this individual's *expertise.* It was this expertise that was used in the knowledge-intensive part and led to the completion of the task in the system. When connecting this individual with another individual seeking advice, the knowledgeable individual is likely able to point to useful sources or help directly drawing from his expertise.

This renders these activities well-suited candidates for determining *task overlap* in a similarity measure. For determining *levels of expertise* they are suited. Seid and Kobsa note that if finding an expert to use him as gateway to or instead of an information record of some kind, it is justified to use the universe of this information records to extract expertise assumptions (Seid, Kobsa 2003, p. 13). In this thesis, those information records are the activities that are observable and that can be associated with "a reason" for performing the activity. As a downside, Seid and Kobsa note that " [...] one may question their basic assumptions on the ground that browsing shows more of the interest of someone rather than expertise." (Seid, Kobsa 2003, p. 8). They were referring to documents and websites that users browse and that have been used as evidence for expertise. However, when operational tasks are used, those are typically not explored but actively used, which remedies this challenge for this thesis' approach. Terveen et al. (2002) also state that activity-related, statistical data such as time spent reading various parts of a document or counts of spreadsheet cell recalculations, i.e. how an IS's functionality is used, can be used to adapt the behavior of ERS. In the context of this thesis, this translates to different methods for determining suitable interaction partners. Activities that are observable, because they are reflected in an IS and that provide some information to infer the individual's goals are also suited to connect individuals. Richter et al. (2008, pp. 18–19) state that in distributed settings, for initial contact as well as for building up trust and the creation of a shared context, IT tools are necessary. The creation of a shared context is easier when the goals of the individuals, possibly only for certain tasks, have an overlap, which is easier to determine using activities that are observable and that convey more context.

5.1.2 Anonymity as an Option

The approach developed in this thesis, aims at connecting physically distributed individuals relying on computer-mediated interactions. Thiedeke (2003, pp. 25–33) noted that in computer-mediated environments, the following four properties hold:

- Anonymity: In purely virtual settings, it is not possible to identify the interaction partner.
- Losing self confining: In purely virtual settings, it is possible to deviate from rules of conduct and to leave social roles as there are no means for sanctioning this behavior.
- Interactivity: Computer-mediated means of communication allow to communicate with more than one individual.
- Optionality: Individuals can choose the topics, forms and intensity of interactions.

Anonymity has been shown in various publications to increase benefits dependant on the setting (see Pinsonneault, Heppel 1997, p. 91 for an overview of research results and a discussion of benefits and downsides). However, those benefits only arise when it is not at all possible to connect the interactions with the individuals that were part of the interaction. For example, if arbitrary usernames may be chosen by an individual, anonymity might be possible. However, if socio-metric information about the user may be used to infer the real-life name this would not be the case. If individuals are not entirely anonymous, there is a possibility for sanctioning inappropriate behavior. One such inappropriate behavior is of social nature. An individual asking for help has the expectation to actually receive help, which in turn limits the optionality of interactions for the queried individual. The choice of the level of anonymity also influences which kind of social fabric might be created through the collective interactions. In distributed networks of practice for example, the individuals know each other, have possibly met each other already, and typically expect to know who they are interacting with. Therefore, anonymity in this case is less expressed or not possible. In electronic networks of practice, that connect many individuals that typically have never and probably will never meet in person, high levels of anonymity are normal. Turning the way of reasoning the other way around, if the mechanisms of an electronic network of practice should be imitated by this thesis' approach, it is necessary to ensure anonymity, however, if establishing a social fabric similar to a distributed network of practice is the goal, anonymous interactions should not be possible. This is why it is a crucial design choice to enable individuals to interact anonymously, non-anonymously or partly anonymously.

The importance of the concept of anonymity is also stressed in related research. With respect to social networking, a somewhat related concept, Richter et al. (2008, p. 20) state that the success of social networking in private usage is strongly related to enabling *awareness* of the other individuals' *current context*. In company settings, they stress, this is however different: There are individuals that like the idea of others knowing what they are doing and others that do not like this idea. Terveen et al. (2002) note that anonymity, especially for ERS, is an important concept for social data mining systems, i.e. systems that try to retrieve information from the interactions of the users, since they gather additional information about the user outside of the original usage context.

Next to the two extremes, anonymous interactions between individuals where neither the asking individual nor the answering individual knows the others' identity, and the other extreme where both interaction partners know who they are dealing with, there are also other possibilities. For example, in their dynamic mailing list solution, Ye et. al. (Ye et al. 2008) expose the sender of a message to all recipients, but it is kept secret who is on the list of recipients. This way the sender does not know who received his message and each receiver only knows that he has received a message from the (named) requestor, but not who else has received the message. This allows each receiver to decide whether or not he wants to answer without the fear of social punishment for not attending to another individual's request. If an individual from the receiver list chooses to answer, the answer is sent to all others on the mailing list, now including the name of the individual providing the answer to foster this individual's reputation as knowledgeable person that is willing to share. A similar approach can be found in (Oldigs-Kerber et al. 2003), where each recipient is also anonymous.

5.1.3 Problems to Address in Expert Determination

Especially in the domain of ERS, there are some issues that remain challenging. This thesis' approach, being a hybrid between ERS and knowledge networks, should try to tackle these challenges. Richter (2008, p. 18) for example notes, that two barriers for ERS are *too little trust* and missing knowledge of *what is happening* in other teams, projects and departments. Overlap in *task context* is one measure that could support fostering trust. Results of empirical studies that researched the correlation of *trust* and *user similarity* give promising indication to support this assumption (Ziegler, Golbeck 2007). In general, however, trust between interaction partners needs to build up over time – an insight from *social capital theory* that was shown to apply to online communities as well (Wasko, Faraj 2005). The lack of knowing what happens in other parts of the company can be addressed by inspecting user's everyday *relevant tasks*.

Next to the pragmatic challenges of building up trust and establishing a better overview of what happens in a company, Seid and Kobsa (2003) analyzed different ERS and determined four gaps between what systems currently offer and what is needed for good ERS.

1. **Source heterogeneity gap**: Since many different kinds of records could be used to infer expertise, the integration of these different systems needs to be addressed.

2. **Expertise recognition methodology gap**: Up until today, the research domain focused more on researching the feasibility of solutions than seeking optimal approaches and techniques.
3. **Expertise analysis support gap**: It is the user who does the selection in the end, therefore, the transparency of the mechanisms why someone is deemed expert should be made more obvious.
4. **Reusability, interoperability and extensibility gap**: Currently, systems focus on one problem but do not pay too much attention on how this can be extended to other systems. Extending to other systems is therefore a critical next step for wider application (Seid, Kobsa 2003, p. 14).

To address the first gap, this thesis' approach should be able to integrate different information sources, which means IS that feature observable functionality that can be interpreted as tasks. The second gap is a general observation about ERS. This thesis sets out to improve the principle working of ERS by combining them with knowledge networks and in this way addresses this gap. Concerning the third gap, a different path is followed. This thesis' approach tries to relieve the individual from the effort of searching for experts. The choice to engage into interaction is shifted from the requestor to the selected (set of) queried individuals, similar to normal interactions in electronic networks of practice. The mediating concept in this case is the similarity of users' tasks. Therefore, still addressing the expertise analysis gap, it is a design choice to indicate the degree of similarity and the number of individuals that are above a certain threshold to the requestor or not. This thesis' approach addresses the fourth gap that asks for wider applicability by proposing a generally applicable concept. It is instantiated taking into account the specific situation as will be shown in the evaluation using case studies.

5.1.4 Design Dimensions

Further acknowledging the relationship to ERS, this thesis' approach adopts some of the ERS features but diverts from the idea of ERS in some important aspects. To better understand the aspects that do relate to ERS, this thesis' approach is next categorized according to two ERS schemas. Along the features within the categorizations, the behavior of this thesis' approach will be described. The first categorization was proposed by Terveen and Hill (2002), the second one by Seid and Kobsa (2003, pp. 10–12) . Terveen and Hill distinguish four dimensions for the design of recommender systems: *Extraction and modeling of preferences*, *distinguishing of roles and communication mechanisms*, *algorithms for recommendation generation* and *human-computer interaction*.

Concerning the design dimension *extraction and modeling of preferences*, this thesis' approach will not rely on direct preferences but relate individuals with a high *contextual overlap* and *expertise* with respect to their task. In this sense, the modeling of preferences is a modeling of previous task executions, i.e. an execution of a task resembles "preferring" this task in the words of general recommender systems or put differently: to assign high utility to this task. Hence, the "preferences", i.e. the previously executed tasks, of seeker and receiver are used for creating the recommendation. They are determined implicitly; the individual has no effort to maintain any explicit information. The extraction is done contingent on the source system. This aspect will be address in more detail in section 5.6.

Concerning *roles and communication mechanisms*, in this thesis' approach, there will be two principle roles, the asking individual and the question receiving individuals. However, the roles may change after each thread of interaction of the individuals. Therefore, for one question an individual can be an asker – a knowledge seeker – and in another one the same individual may receive a request that he answers, he is an expert – a knowledge bearer – in this case. Also within one interaction thread, the question-answer paradigm may fluently change to a more discussion-like interaction, which is common in knowledge networks. In terms of communication, this thesis' approach will allow communication in an one-to-one interaction between asking individual and answering individual, but store the communication openly in a knowledge network-like structure and in this sense is one-to-many – a hybrid solution as also proposed in (Ye et al. 2008) for example.

The *algorithm for generating recommendations* will operate on the tasks that can be extracted from the source system. It should apply a suitable similarity measure to determine the contextualized task overlap of individuals. Additionally, the recommendation may use the number of times that an individual has executed a task to use it as an evidence for this individual's expertness. In essence, there will be two ingredients for the recommendation: the *evidence of expertness* and the *contextual overlap* of asker and potential answerer.

The final dimension in Terveen and Hill's (2002) classification is *human-computer interaction*. To this respect, this thesis' approach is different. Terveen and Hill see this dimension's interpretation as the representation of a list of possible interaction partners. In this thesis' approach, the notification of possible "experts" is done behind the scenes. The asking individual sends a request, while the notification of suitable receiving individuals is then executed automatically and possibly anonymously. Therefore, in this thesis' approach there is no human-computer interaction with respect to Terveen and Hill's interpretation.

The second classification schema for ERS that will be used for structuring this thesis' approach and for informing its design, is the one by Seid and Kobsa (2003, pp. 10–12) . They identify seven domain factors that they use to classify ERS:

1. **Basis for expertise recognition**: Identifies the potential sources of expertise.
2. **Expertise indicator extraction**: Differentiates between domain-specific and domain-independent extraction of expertise indicators.

3. **Expertise models**: Related to a suitable meta-description of an individual's expertise. It can be created directly (with manual effort) or automatically. In the latter case, the model always carries a degree of uncertainty in how well it really reflects the expertise of the individual. In addition, the question arises how and when and how it is compiled. This can be done during query-time or prior to the query, as well as distributed or centralized.
4. **Query mechanism**: Distinguishes how the query for an expert is created. This can be either explicitly by the user or implicitly due to context factors.
5. **Matching operation**: This factor addresses how the query is matched with the expert profile, i.e. which algorithmic approach is used. Possibly, suitable inference mechanisms are considered here as well.
6. **Output presentation**: This factor addresses how the result of the expert query is presented to the user, e.g. if it is presented as a ranked list or unordered list.
7. **Adaptation and learning operation**: This factor addresses the ERS's ability to become more and more adapted to the user by learning from explicit or implicit user feedback.

This thesis' approach will build upon the use of task overlap to find suitable interaction partners, which are the equivalent of experts in ERS. Therefore, the basis of "expertise" recognition is the task sequences that are extracted from logs of IS (*domain factor 1*).

The domain is determined by the tasks that are deemed relevant for the overlap determination and are from a conceptual point of view domain-independent, but, depending on the application scenario, could be restricted to only certain tasks that are then domain-dependant (*domain factor 2*).

This thesis' approach will not build up an expertise model in the classical sense. Instead, the tasks that are executed by an individual, their sequence and numbers of repetition are retrieved to give indication of expertise as well as contextual overlap. The task sequences are available automatically in the form of log files of operational systems, but need to be extracted and transferred to where the computation is done. The computational use of the tasks and their sequence may however be ad-hoc to a certain degree. With respect to data acquisition, Seid and Kobsa (2003, p. 16) also discuss the different implementation's balance between up-to-dateness and performance. The query-time generated expertise models face a performance problem, but are by-definition up-to-date as they query the underlying information base during the query. The apriori generated models have in general less performance impact but may not be fully up-to-date. This thesis' general approach is independent of this, but acknowledges that a hybrid approach may be beneficial. The later implementation of the approach in concrete systems will have to attend to this challenge (*domain factor 3*).

The query for suitable interaction partners is triggered by the user's explicit need and explicit statement of a knowledge seeking request. However, the determination itself is then automated, i.e. the requesting individual is not querying expertise as such, but rather taking the context of the current user and the request itself, suitable interaction partners are determined automatically (*domain factor 4*).

The matching mechanism capitalizes on two things. A similarity of the requesting individual's task or task sequence with those of possible interaction partners to reach a satisfactory degree of contextual overlap and the number of times a possible interaction partner has executed a task to get some indication for expertness in executing the task (*domain factor 5*).

This thesis' approach connects interaction partners in a non-obstructive way for both the searching party, who just asks without needing to search for an expert and hence with reduced mental effort and, on the other hand, for the receiver who is notified about a request for support, but is free to choose if he wants to contribute or not (see also 5.1.2). Therefore, in a stricter sense, there is no result of expert determination that should be shown. In a broader sense, the use of knowledge networks is this thesis' form of representation, as useful connections between interaction partners will result in the creation of a discussion that can be shown publicly, similar to how it is in knowledge networks (*domain factor 6*).

Since task (sequences) are utilized as "reference unit" for the determination of expertise, there are some general ways, how adaptation to the user and learning of preferences is possible. Possible mechanisms are the limitation of the backward-facing window of tasks that indicates which tasks should influence the determination of contextual overlap and expertness of possible interaction partners. This way, recent interactions enjoy more importance. Additionally, the use of the measure of contextual overlap can be adapted. Further, it might be the case that a user starts to become reluctant to answer questions concerning a certain topic and might indicate that it is not relevant for the person any longer. The same might be true for persons, with whom the individual does not like to interact (anymore). The selection of suitable means is certainly dependant on the use case and hence implementation-specific, but this thesis' approach should cater for it (domain factor 7).

5.1.5 Embeddedness

Section 4.3.4 discussed the requirement of IT-supported KM tools to integrate into the operational processes. In its most strict sense, this requirement implies that the support function, in the context of this thesis the support of connecting suitable interaction partners, should be integrated not only as part of the (business) process but ideally also into the same tool that the knowledge worker uses for his tasks. In the field of ERS, Seid and Kobsa state that "[...] since expert-seeking arises during day-to-day work, it is beneficial if expert finding systems are embedded in the day-to-day problem solving and information search environments" (Seid, Kobsa 2003, p. 13) which underlines the importance of integration into operational IS. In the same line of argumentation, Qureshi et al. (2006) state that when individuals rely on the support of others they appreciate immediate help that is focused to their particular needs. If a supportive facility is integrated into the operational IS, immediate help is easier and faster to get, and the contextual information of the current work situation can be exploited to better determine the current need and hence suitable forms of help. Therefore, this thesis' approach should explicitly address the need to integrate into operational IS or should give guidance for actual implications of the concept to this ends.

5.1.6 Reducing Information Overload

Section 4.2.11 concerned itself with motivation as a success factor for KM projects in general. It showed that one form of motivation is to create benefits for the individual such as increased efficiency and effectivity of the individual's work. The opposite of this, the perceived decrease in efficiency and effectivity due to the "extra burden" of KM activities is therefore a relevant barrier. Consequently, the effort and necessary time for KM activities should be kept low. The importance of acknowledging the individual's value of time was stressed by a survey conducted by a consultancy among 413 individuals. The results are shown in Figure 5-3.

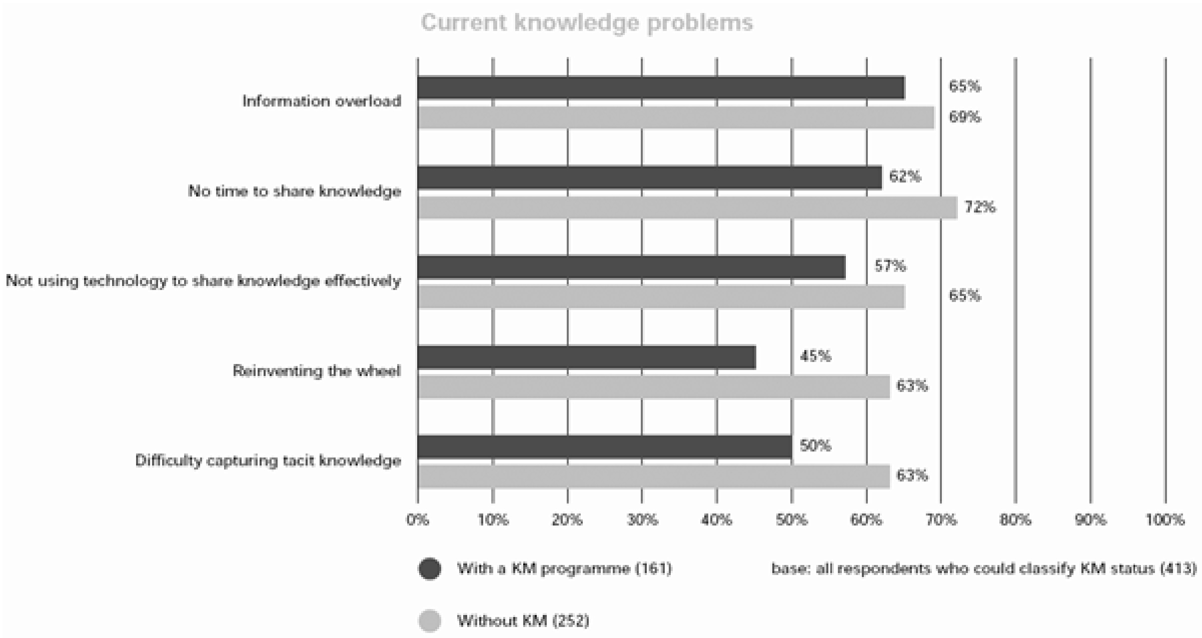


Figure 5-3 ***Reported problems with Knowledge Management***
Source: Adapted from (KPMG 2000)

The two most often mentioned problems for KM programs are a *lack of time* to share knowledge and a feeling of getting too much information, i.e. *information overload*. For organizations that have no KM initiative in place yet, the problem of too little time was mentioned in a considerable 72 % of the responses and information overload, being mentioned as a problem in 69 % cases, follows shortly after. The Frauenhofer IAO institute together with Deutsche Bank conducted a similar survey and came to comparable results: The feeling of being short in time was by far the most often mentioned problem that individuals see in KM activities (Frauenhofer IAO und Deutsche Bank 1999). This is in line with general thoughts of Simon who states that " [..] it is human attention, not information, that is the scarce factor" (Simon 2002, p. 614). Therefore, this thesis' approach should try to reduce information overload and should try to improve the time necessary to participate in KM activities.

In order to reduce the time necessary for KM activities and more specifically sharing knowledge by means of direct interaction between individuals, it is necessary to understand where potentials for improvement lie. Ye et al. (2008) have analyzed the consumption of time that happens in a collective of individuals when situated knowledge requests are issued. As depicted in Figure 5-4, they distinguish between the asker and the recipients. The recipients are further divided into those that simply ignore a request, those that skim the question, and those that try to answer the question.

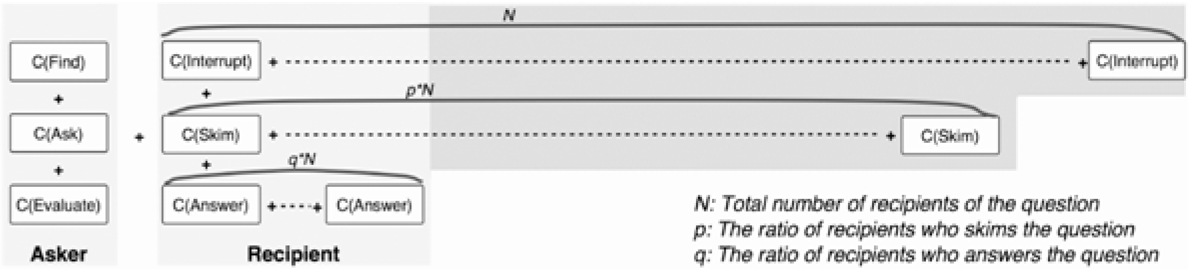


Figure 5-4 ***Schematic calculation of the cost of collective attention***
Source: ***(Ye et al. 2008, p. 170)***

The asker needs to invest time and effort into finding suitable individuals to ask, to properly formulate a question, and to evaluate the recipients answers with respect to their correctness and applicability. All recipients of the question are distracted from their work when receiving the request. This short interruption may inflict only small distraction but the number of possibly affected individuals may be high. To evaluate the question and determine if the individual may answer the question, at least skimming over the question is necessary, which consumes additional time. If the recipient is able to and willing to answer, there is additional effort for formulating an answer. In sum, the effort can abstractly be formulated in the following equation:

$$Effort = 1 * Asker_{find} + 1 * Asker_{ask} + q * N * Asker_{evaluate} + N * Recipient_{interupt} + p * N * Recipient_{skim} + q * N * Recipient_{answer}$$

Ye et al. (2008, p. 171) indicate that the cost-benefit-ratio of the described setting can be improved in two ways. The consumed attention for each interaction can be reduced and/or the result of the interactions, i.e. the quality of the answers can be increased, which they denote with another variable – the *success rate r* – that indicates how many of the posed questions are answered satisfactory. In their article, they describe a solution that improves the cost-benefit ratio by forwarding questions only to those recipients that are likely to appreciate receiving the question and answer it. They propose to use information stored in YPS or automated expertise extraction systems similar to ERS for determining expertise and hence who should receive questions. Their general approach appears suitable for this thesis' KM approach as well and will be adapt to the core concept, the overlap of task contexts, to reduce information overload.

5.2 Theoretical Basis

The design of this thesis' approach is informed by the results of chapter 3 that determined the downsides of currently employed solutions for fostering knowledge exchange and the factors to keep in mind for successful KM approaches as determined through literature analysis in chapter 4. However, another source of valuable information is used for creating an adequate design of the approach. Theories that describe contribution to knowledge networks and empirical evidence that was drawn using these models form a suitable basis on which to build artifacts such as this thesis approach. Taking the results from theory to inform the design of an artifact is a cornerstone of design science and ensures rigor (Hevner et al. 2004, p. 88). It also may lead to finding non-intuitive designs that capitalize on theories of cause and effect, also referred to as *theory-driven design* (Card 1989; Briggs 2006).

The models analyzed in the following comprise one model describing *contribution behavior* based on psychological theories and one strand of empirical results based on *social capital theory*. In the following, both theoretical foundations and their influence on the later design are detailed.

5.2.1 Models Based on Social Capital Theory

Properly supporting computer-mediated interactions and knowledge transfer in particular requires to understand the social mechanisms that govern those interactions. Ackermann and Halverson believe that *Social Capital Theory (SCT)* is a driving force in this endeavor by stressing that "[...] significant benefits will accrue only with understanding the need for social capital and incorporating its dimensions into all types of knowledge management technologies." (Ackerman, Halverson 2004, p. 16). To understand how to incorporate SCT, there have been some empirical studies out of which two especially suitable ones are inspected in the following.

Wasko and Faraj (2005) investigated what motivates individuals to contribute in an electronic network of practice (see section 3.4.6.3) where there is no obvious direct benefit and free-riders benefit equally to everyone else. They find that those mechanisms that have previously been reported to positively influence knowledge sharing such as *strong ties*, *co-location*, a history of *prior relationship* and *demographic* and *status similarity* are all not apparent in electronic networks of practice due to the relative anonymity among the participants. They propose to apply SCT (see for example Lin (2001) for an overview) to understand the behavior of the participants, more specifically the adaptation of SCT to knowledge contribution (Nahapiet, Ghoshal 1998). According to Nahapiet and Ghoshal's model, exchange of knowledge is facilitated, if the interaction partners are *individually motivated*, there are structural links between them (*structural capital*), they are able to understand and apply the knowledge (*cognitive capital*) and the relationships between them are positive and strong (*relational capital*). Wasko and Faraj (2005) operationalize those four influencing factors as shown in Figure 5-5. They applied their model in an electronic network of practice where they could obtain 173 responses to a survey and used those to test their hypothesis.

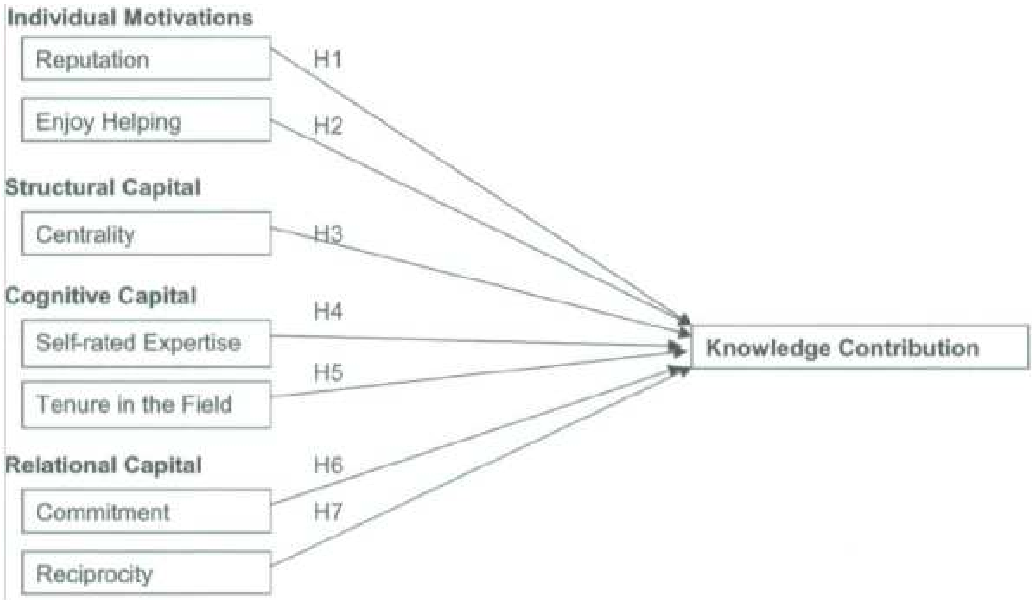


Figure 5-5 ***Wasko and Faraj's model for knowledge contribution in electronic networks of practice***
Source: (Wasko, Faraj 2005, p. 40)

The results of the statistical analysis provided the following insights. *Reputation* is a strong predictor for knowledge contribution, while *enjoyment* had no significant influence. Wasko and Faraj reason that the relative lack of anonymity in the network that they researched could be a reason for much more pronounced extrinsic motivation than intrinsic. Further, they found that structural capital, i.e. *centrality*, is the most significant predictor for knowledge contribution. The cognitive capital influenced knowledge contribution differently than expected. While *tenure in the field*, i.e. the experience of an individual, significantly influences knowledge contribution, *self-rated expertise* does not. Wasko and Faraj reason that this difference to previous studies' results may be due to how the expertise was measured. Interestingly, relational capital could not be shown to influence knowledge contribution, so neither commitment to the network nor expectations of direct reciprocity seemed to influence the willingness to contribute.

Also building upon the SCT model for knowledge contribution (Nahapiet, Ghoshal 1998), Law and Chang (2008) investigated which IT artifacts contribute to the build-up of different aspects of SCT that then, according to the previous study, should lead to knowledge contribution. By observing technology features of various online communities and additional literature review and participation in discussions in communities, they found four IT artifacts to be relevant for their model: identity profiling, sub-community building, feedback mechanism, and regulatory practice. *Identity profiling* refers to the ability of presenting oneself to other community members. *Sub-community building* refers to the ability to engage in and manage membership of smaller groups within the larger community. *Feedback mechanism* refers to peer reviews, evaluations, and ratings, i.e. feedback channels between individuals after interactions. *Regulatory practice* refers to any rules, procedures, and guidelines that are supplied and enforced by the community platform / provider. These IT artifacts are expected to influence different aspects of SCT as shown in Law and Chang's (Law, Chang 2008) research model depicted in Figure 5-6.

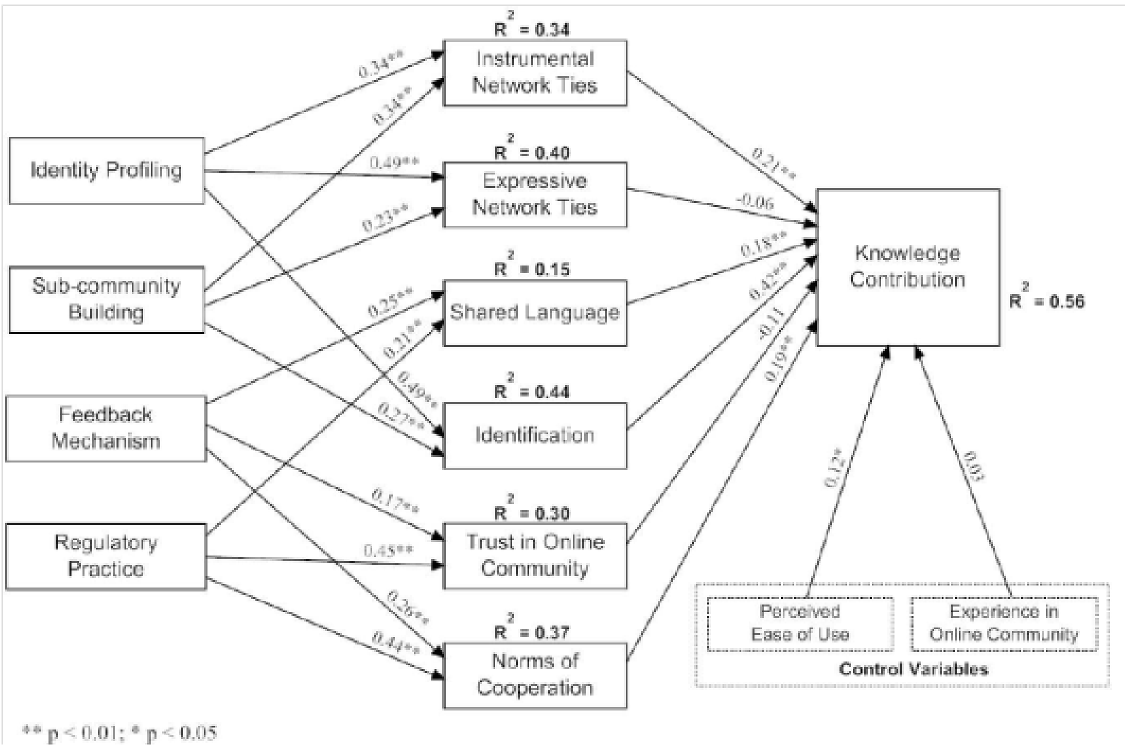


Figure 5-6 ***Law and Chang's model for knowledge contribution in electronic networks of practice***
Source: (Law, Chang 2008, p. 15)

Their statistical analysis of survey results from 253 respondents across communities of common interest and information exchange as well as samples of students and alumni from their university, revealed that all hypothesis of Figure 5-6 could be supported except, the influence of expressive network ties on knowledge contribution and the influence of trust on knowledge contribution. Identification could be shown as most influential factor.

An interesting result is that the two studies of Law and Chang (2008) and Wasko and Faraj (2005) both build upon the same theory, namely SCT for knowledge contribution (Nahapiet, Ghoshal 1998), but, next to different scopes, differ in the (relative) influence they determine for similar constructs. For example, *identification*, part of relational social capital, was shown to be most influential in Law and Chang’s study while relational capital was shown to be not influencing knowledge contribution in Wasko and Faraj’s results. Since in the two studies, different kinds of communities were inspected – Wasko and Faraj looked at a network of a professional association, while Law and Chang looked at public online communities – social capital might, as Law and Chang indicate, be context-specific to the network itself (Law, Chang 2008, p. 17).

Taking the results of these two empirical studies into account, SCT’s influence on contribution behavior seems to hold explanatory power although settings were inspected where individuals were not co-located, individuals did not know each other, and in the case of Law and Chang’s (2008) study, did not have any affiliation at all. This suggests that for the design of this thesis’ approach one may draw upon SCT, as it also focuses on distributed settings. Neglecting the hypotheses that could not be proven in either studies and focusing on the ones that could be proven, this thesis’ approach should hence cater for the following.

It should try to *foster reputation* as a strong motivator especially in professional networks, where enjoyment of help is less crucial. Further, the more central users are with respect to (instrumental) network ties, the more likely they are to contribute knowledge, and hence this thesis’ approach should try to *increase centrality* at least indirectly by increasing the “offers” for communication ties. was shown to be important (Wasko, Faraj 2005), but cannot be influenced by a mediation method, but must rather be supported by the company, hence it is an implementation-specific feature. Shared language was shown to influence knowledge exchange and hence this thesis’ approach should try to connect individuals that may converse in a common language, for example by connecting people that *perform similar tasks*. Identification with the network and its members was shown to influence knowledge exchange in some cases (Law, Chang 2008, p. 17) but not in others (Wasko, Faraj 2005). As there seemed to be no negative effect for increasing identification, this thesis’ approach should support the build-up of identification with other network members. This might be possible, when individuals interact with other, *similar individuals*, as similarity in personal traits is a strong mechanism for identification with others and social groups in particular (Tajfel 1986; Turner et al. 1987). Norms of cooperation are again an implementation-specific factor, which this thesis' design cannot integrate into its general design while implementations of the general design certainly have to take care of this aspect. Relating to the IT artifacts that Law and Chang (2008, p. 17) identified as supportive for the formation of social capital – identity profiling, sub-community building, feedback mechanism and regulatory practice – all could be shown to be relevant and hence this thesis' approach should incorporate mechanisms for those respective artifacts.

5.2.2 Model of Contribution Behavior Based on Problem Solving Theory

Olivera et al. (2008) developed a model describing how and why people contribute in distributed organizations through IT-mediated means. They use the individual that receives a request as unit of investigation pointing out what prompts this individual to formulate a response to a request for help. They argue that to understand the contribution behavior, two strands of theories have to be combined: theories of *problem solving* (Newell et al. 1972) and *cognitive motivation* theories (Kanfer 1990). The first one describes how human beings address problem solving, distinguishing three activities: *awareness* of problem, *searching and matching* expertise or previous experiences and *formulation and delivery* of an answer. The second, cognitive motivation theory, describes which factors are necessary to have an individual bear the effort of the mental steps in creating a response to another individual's request. They interpret the three problem solving activities as mediators between contextual factors and the final contribution.

The first activity is *awareness* which is "[..] a cognitive activity through which a person recognizes an opportunity to contribute" (Olivera et al. 2008, p. 26). The result of this activity is a mental representation of the problem stated in the request and an individual level of motivation. The second activity is *searching and matching* in which "[..] individuals determine whether and how the knowledge domain of the help request matches their own personal knowledge [..] and may require an individual to search both his or her internal memory and external memory aids [..]" (Olivera et al. 2008, p. 26). As this activity inflicts cognitive effort and hence cost for the potential respondent and because longer searches inflict higher costs, the process of contributing may be interrupted at this step, if matches cannot be found. The last activity *formulation and delivery* is an "[..] activity through which the contribution is articulated and communicated" (Olivera et al. 2008, p. 27), where formulation encompasses the determination of what should be communicated and delivery determines the communication means.

Seeing that the three mediating activities induce "cost" on the potential respondent, there need to be motivational forces that overcompensate the cost, otherwise there will be no contribution (see also section 4.2.11). Consequently, in Olivera et al.'s eyes "[..] motivation is both a consequence and an antecedent to the mediating mechanisms." (Olivera et al. 2008, p. 28). Drawing upon existing literature on altruistic and pro-social behavior in organizations, they find three classes of motivations that are especially important: Self-enhancement, exchange motivation and instrumental motivation. Motivations of *self-enhancement* prompt an individual to action because he likes to express his expertise (Wasko, Faraj 2000, p. 166) and live up to his self-identity (Constant et al. 1996, p. 122), likes to learn through helping (Wasko, Faraj 2000, p. 166) or finds enjoyment in helping others (Wasko, Faraj 2000, p. 166; Constant et al. 1996, p. 126). *Exchange motivations* refer to the expectation of reciprocity (Blau 1986). It is a natural human desire to balance giving and taking. There may be direct reciprocity, i.e. an individual helping another individual expects to have "the favor returned" by this individual in the future. In addition, there may be generalized reciprocity (Wasko, Faraj 2000, p. 158), i.e. an individual helping another (often anonymous other) individual expects likewise treatment by someone from the social group to which that individual belonged but not necessarily the exact same individual. The expectation of reciprocity was shown to be especially relevant when there are weak norms of sharing (Kankanhalli et al. 2005, p. 26). *Instrumental motivation* refers to receiving material (e.g. Markus 2001, p. 83) or immaterial reward such as recognition (O'Dell, Grayson 1998, pp. 163,168).

Utilizing the two theoretical bases of motivation and problem solving, Olivera et al. (2008) complement their model by taking into account ten contextual constructs that should interact with the mediating mechanisms. The entire model is depicted in Figure 5-7. Reasoning about the interactions of the two theories and the contextual factors, they deduce seventeen propositions (Olivera et al. 2008, pp. 29–35)– eight for the mediating activity *awareness*, four for the mediating activity *searching and matching* and five for the mediating activity *formulation and delivery*. In the following, those propositions are summarized and incorporated or adequately adapted to fit the design of this thesis approach:

1. Requests from senders of higher status than the recipient will generate a higher motivation to engage in searching and matching (characteristic: *sender status*; *instrumental motivation* as trigger)
2. Requests from senders with affiliations with which the recipient has had prior positive exchanges will generate a higher motivation to engage in searching and matching (characteristic: *sender affiliation*; *exchange motivation* as trigger)
3. Requests in a topic domain in which the recipient has expertise that is central to his or her identity will generate a higher motivation to engage in searching and matching (characteristic: *request domain*; *self-enhancement* as trigger)
4. Requests that are specific will generate lower anticipated cost and in turn higher motivation to engage in searching and matching (characteristic: *request specificity*; *effort-benefit relation* as trigger)

5. Requests that are concrete will generate lower anticipated costs and in turn higher motivation to engage in searching and matching (characteristic: *request concreteness*; *effort-benefit relation* as trigger)
6. Requests that are specific and abstract will generate lower anticipated costs than requests that are general and concrete (characteristic: *request concreteness* and *specificity*; higher *sensitivity to immediate costs* as trigger)
7. Requests communicated through media that are high in social presence will generate higher motivation to engage in searching and matching (characteristic: *social presence*; communication of *urgency* and *socio-emotional concern* as trigger)
8. General requests communicated through synchronous media will generate higher motivation to engage (characteristic: *synchronicity* and *request concreteness*; *media richness theory* (Daft, Lengel 1984) and *cost-benefit relation* as trigger)
9. The searching and matching activity may itself generate self-enhancement motivations and hence motivation to engage in formulation and delivery (characteristic: none; *joy of solving problems* (see above paragraphs) as trigger)
10. If and when searching and matching moves from internal to external memory systems (i.e. it becomes more costly), motivation to engage in formulation and delivery will decrease (characteristic: none; *cost-benefit relation* as trigger)
11. Searching and matching for specific requests will generate a higher level of motivation to engage in formulation and delivery (characteristic: *request specificity*; *cost-benefit relation* as trigger)
12. Using effective search, indexing and retrieval technologies will increase the motivation to engage in formulation and delivery (characteristic: *quality of search support*; *cost-benefit relation* as trigger)
13. Formulation and delivery of a contribution are more likely to be completed, when the request is concrete (characteristic: *request concreteness*; *cost-benefit relation* as trigger)
14. A high investment of time and effort in searching and matching will increase the likelihood of completing the formulation and delivery (characteristic: none; *escalation of commitment* (Staw 1981) as trigger).
15. The likelihood of completing formulation and delivery decreases with anticipated follow-up requests as result of the contribution (characteristic: none; *cost-benefit relation* as trigger)
16. The greater the number and the more accessible communication channels are, the greater the likelihood that individuals who use them will complete the contribution (characteristic: *Communication channel diversity* and *accessibility*; *media richness theory* (Daft, Lengel 1984) and *cost-benefit relation* as trigger)

17. Reduced cost of formulation and delivery due to high-quality authoring tools will increase the likelihood of completing a contribution (characteristic: *quality of authoring tool*; *cost-benefit relation* as trigger)

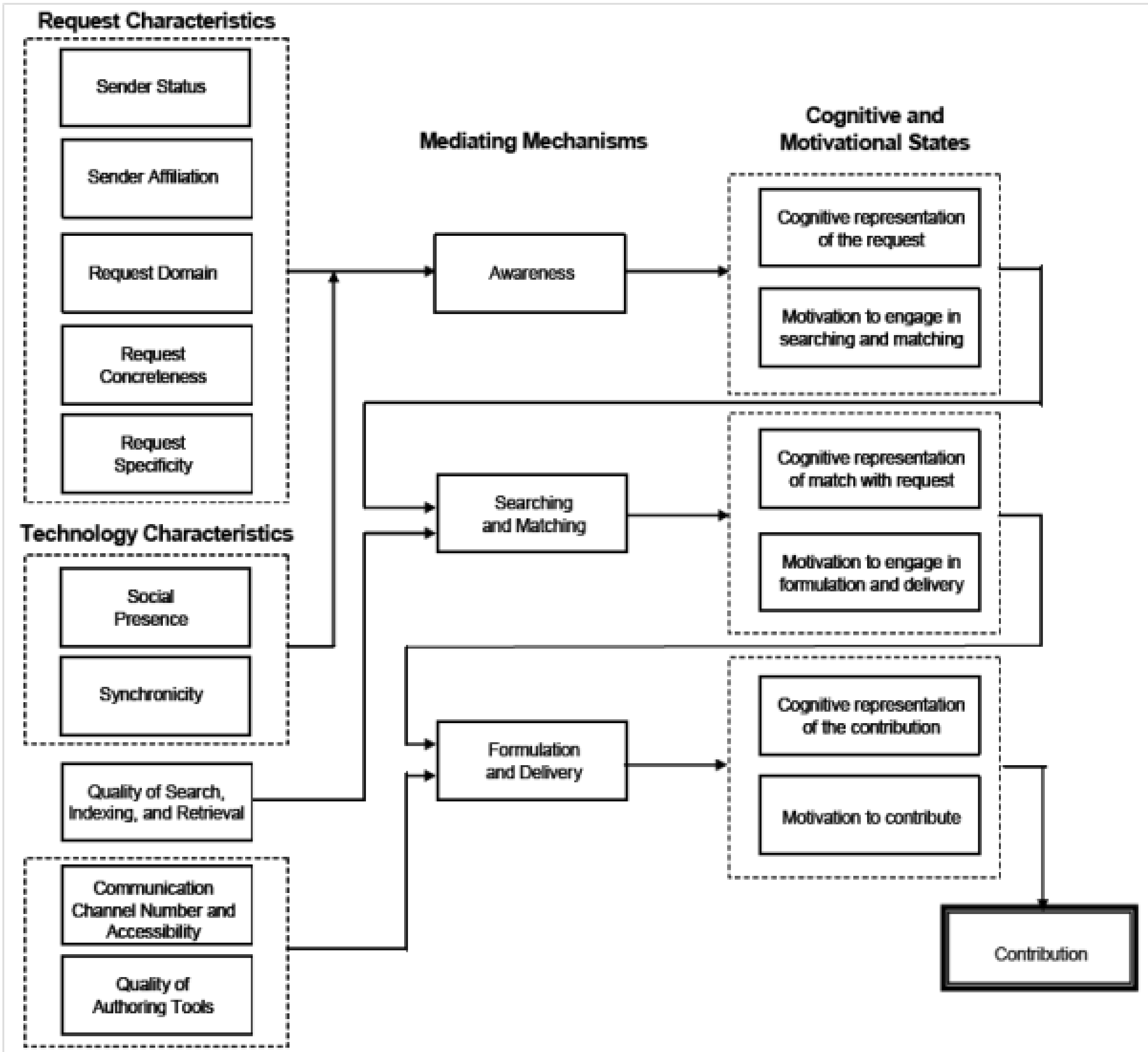


Figure 5-7 ***Olivera et al.'s model of contribution behavior***
Source: (Olivera et al. 2008, p. 27)

Olivera et al.'s model is deduced from different sociological and psychological theories, that each have been proven to hold but there is no empirical validation of the proposed effects of their combination. Therefore, the approach developed in the following will incorporate the possible positive effects of all of the above hypothesis.

5.3 Overview of Approach

Having analyzed contemporary approaches and their problems (chapter 3), having collected success factors for KM initiatives (chapter 4.2) as well as for KM systems (chapter 4.3) and taking into account the preliminary considerations above (chapter 5.1), in this chapter the design of this thesis' approach is highlighted. All factors that have been identified to influence the design of this thesis' approach are summarized and visualized in Figure 5-8.

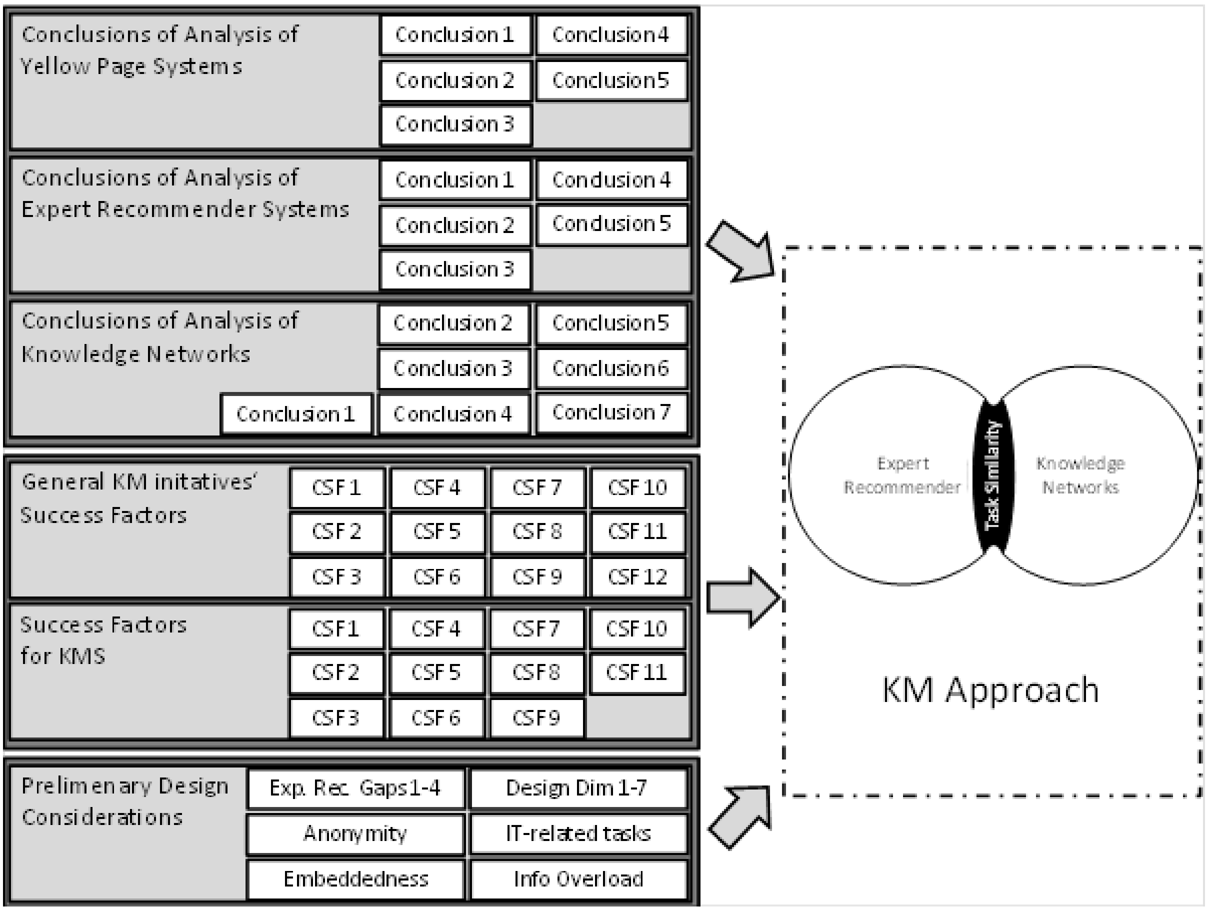


Figure 5-8 ***Influencing factors for this thesis' KM approach's design***
Source: Own illustration

The approach developed in the course of this thesis builds upon the combination of ERS and knowledge networks using task similarity as "glue" between the two approaches. The use of any ERS requires suitable data to operate on and this will be no different for this thesis' adapted approach. The approach capitalizes on the support of a computerized system, hence application logic will perform tasks like determination of the contextual overlap of individuals and management of the communication flows and this way mediates the capabilities of ERS to means for knowledge exchange in knowledge networks. These supporting facilities are hidden from the benefiting individuals. The effects of them however foster knowledge exchange in what looks like a knowledge network to the benefiting individuals. Hence, there are elements that the final benefiters of the system will "see" and interact with – without explicitly knowing about the supporting facility and kinds of data underneath. This is why the approach is described by structuring it according to the three layers as depicted in Figure 5-9. In the following, this thesis' KM approach is introduced on each layer in a black-box view before details are explained for each of the building blocks in the following subchapters.

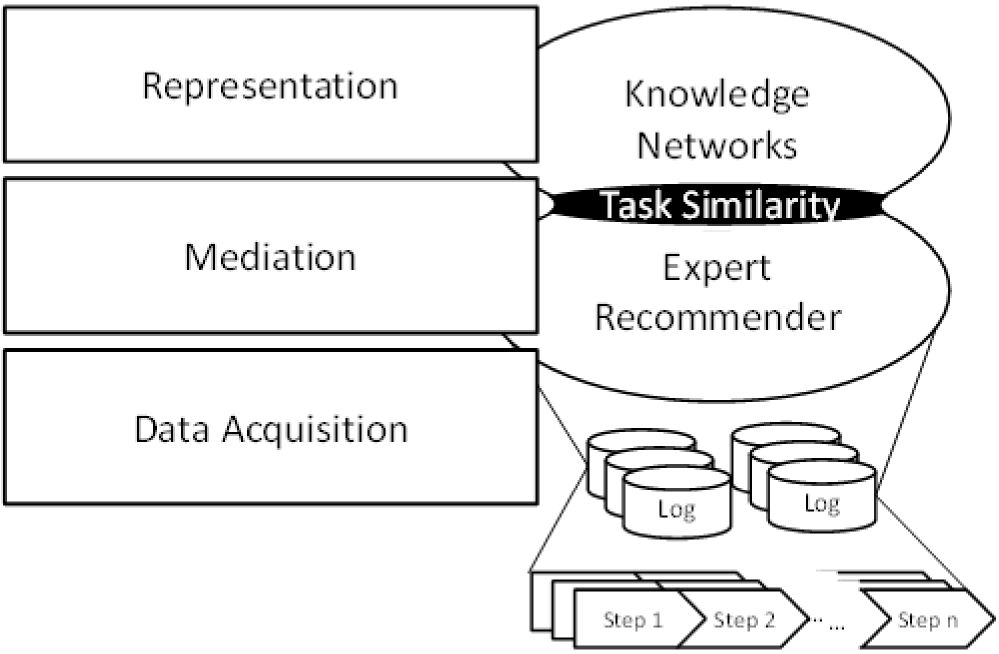


Figure 5-9 ***The three layers of this thesis' KM approach***
Source: Own illustration

5.3.1 Representation

The first layer of this thesis' KM approach – the *Representation Layer* – concerns itself with fostering the knowledge exchange by suitably interacting with the users or more precisely by suitably moderating the interactions among the users. This layer capitalizes on the support of the other two layers. This layer will have to address various challenges and design considerations. In the following, these challenges are described by referring to the source of challenge in brackets after each statement. Additionally, Figure 5-10 visualizes which influencing factors determine the design of the building block *Representation*.

The interaction with other users through the mediating system proposed in the approach should be **integrated into the system, where the need for querying other users originates**, i.e. without media break and with **awareness of the current work context** of the user (Analysis of YPS: Conclusion 4 and Analysis of ERS: Conclusion 4 and KMS CSF 3 and preliminary design consideration "embeddedness" and ERS design dimension 4 and general KM CSF 8). Further, the representation should be created in a way that **interactions between individuals also contribute to the learning of others** that are not or only peripherally connected to the interaction (Analysis of YPS: Conclusion 5 and Analysis of ERS: Conclusion 5 and Analysis of Knowledge Networks; Conclusion 4). Representation should be **structured in a way that is natural for the users** (general KM CSF 6 and KMS CSF 5 and KMS CSF 6) and that allows **managing the content over time** (KMS CSF 4). As a next design consideration, the representation layer must support individuals in **finding already existing possibilities** for knowledge exchange (Analysis of Expert Networks: Conclusion 1). If there are possibilities for knowledge exchange but they have become ineffective due to increased size and diversity of topics, another functionality of the representation layer should support the **formation of smaller groups** (Analysis of Expert Network: Conclusion 3). On the other hand, if there are no suitable possibilities for knowledge exchange the representation layer should be able to **support the formation** of them (Analysis of Expert Networks: Conclusion 2). The representation for the receiver of interaction requests should be designed in a way that **increases the likelihood of contribution** (Analysis of Expert Networks: Conclusion 6, ERS design dimension 7) and keeps the collective information flood at a minimum for all participants (preliminary design consideration *information overload*). In this layer, there will also be design decisions that depend on how the approach should be used, as the approach should be applicable and reusable in multiple contexts (ERS gap 4). Taking into account the organizational culture (general KM CSF 3), **different levels of anonymity** may be appropriate (preliminary design consideration *anonymity*) and **different KM roles** may be necessary (general KM CSF 4) that may have different rights (KMS CSF 7). Furthermore, **different ways of incentifying** expected behavior may be appropriate (general KM CSF 9) some of them possibly relying on **proper activity measurement** for which the layer should also provide means (general KM CSF 5). These criteria may lead to different instantiations of the approach, e.g. the interactions may resemble an electronic network of practice or they may resemble a distributed network of practice and hence there are different, interpretations of what is the right system contingent on the use case (KMS CSF 2).

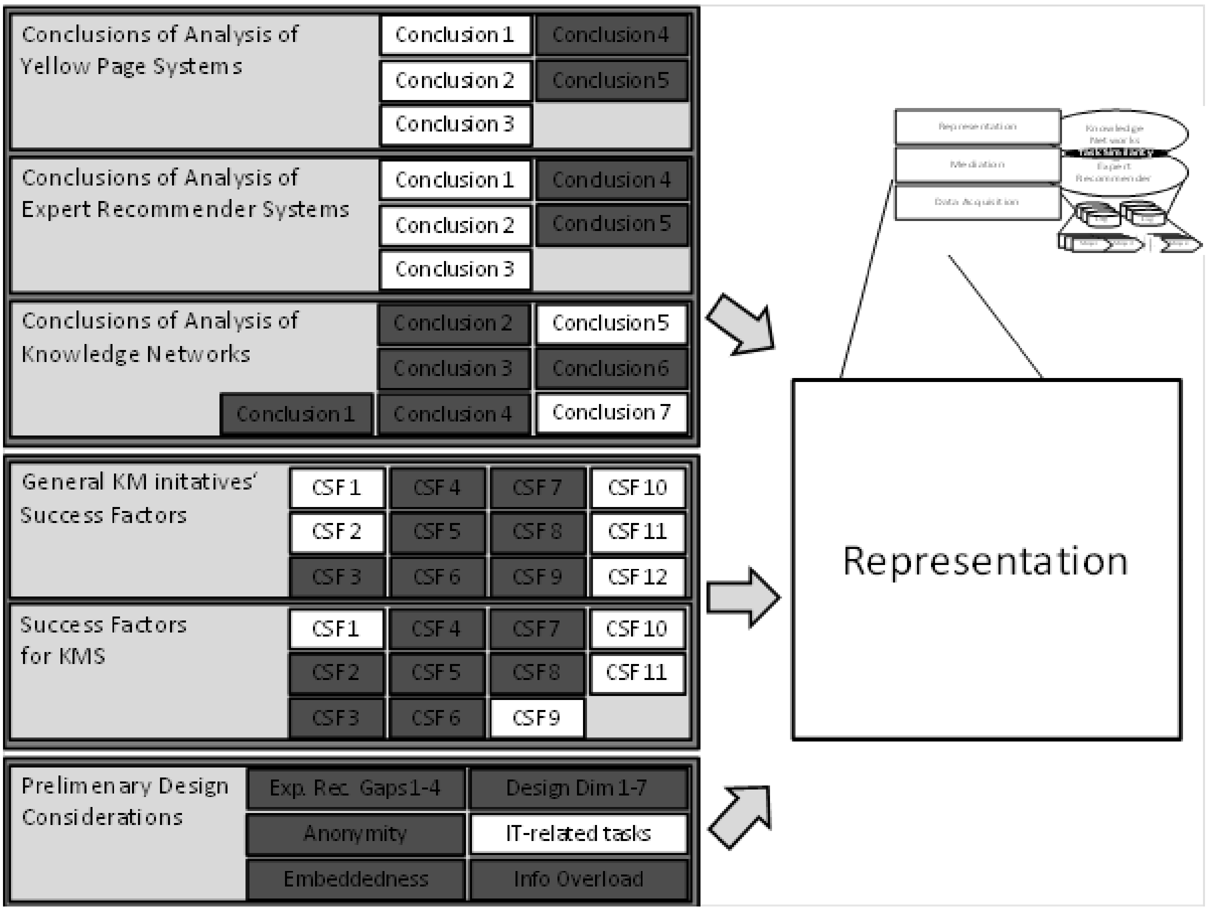


Figure 5-10 ***Influencing factors for the building block Representation***
Source: Own illustration

5.3.2 Mediation

The *Mediation Layer* operates on the data obtained from the data acquisition level to offer any necessary support for knowledge exchange that is later delivered to the users through the representation layer. The provided functionality strongly resembles the inner workings of ERS but deviates from it in important aspects. Its main task lies in matching the requesting individuals with potential interaction partners in an appropriate way. Within this layer, a number of challenges and design considerations have to be taken into account. In the following, those are described by referring to the source of challenge in brackets after each statement. Additionally, Figure 5-11 visualizes which influencing factors determine the design of the building block *Mediation*.

Firstly, this layer will have to provide for the ability to **operate on explicit or implicit models of users** with respect to their competencies and context. It will be necessary to adapt the models whenever changes are propagated from the data acquisition layer (Analysis of YPS: Conclusion 3). Further, a choice of how to find a match between a knowledge seeker and potential interaction partners should be made. It is not always beneficial to find the **most knowledgeable person** for a task, **but contextual overlap is relevant as well**. The deliberate design choice how to fine-tune the determination with respect to these two factors will depend on the instantiation of the approach in a concrete use case (Analysis of ERS: Conclusion 2 and general KM CSF 6 and general KM CSF 9 and KMS CSF 6 and ERS gap 2 and ERS gap 4 and ERS design dimension 3 and ERS design dimension 7). Likewise, the **context of the requesting individual should be used** to obtain a set of candidates that suits the current needs (Analysis of YPS: Conclusion 4 and Analysis of ERS: Conclusion 3 and Analysis of ERS: Conclusion 4 and Analysis of Knowledge Networks: Conclusion 6 and KMS CSF 3 and ERS design dimension 5). The determination of interaction partners in the Mediation layer should be done **without effort for the user** (Analysis of Expert Networks; Conclusion 5) and should try to **limit the consumed collective attention** for interactions (preliminary design consideration *Information Overload*).

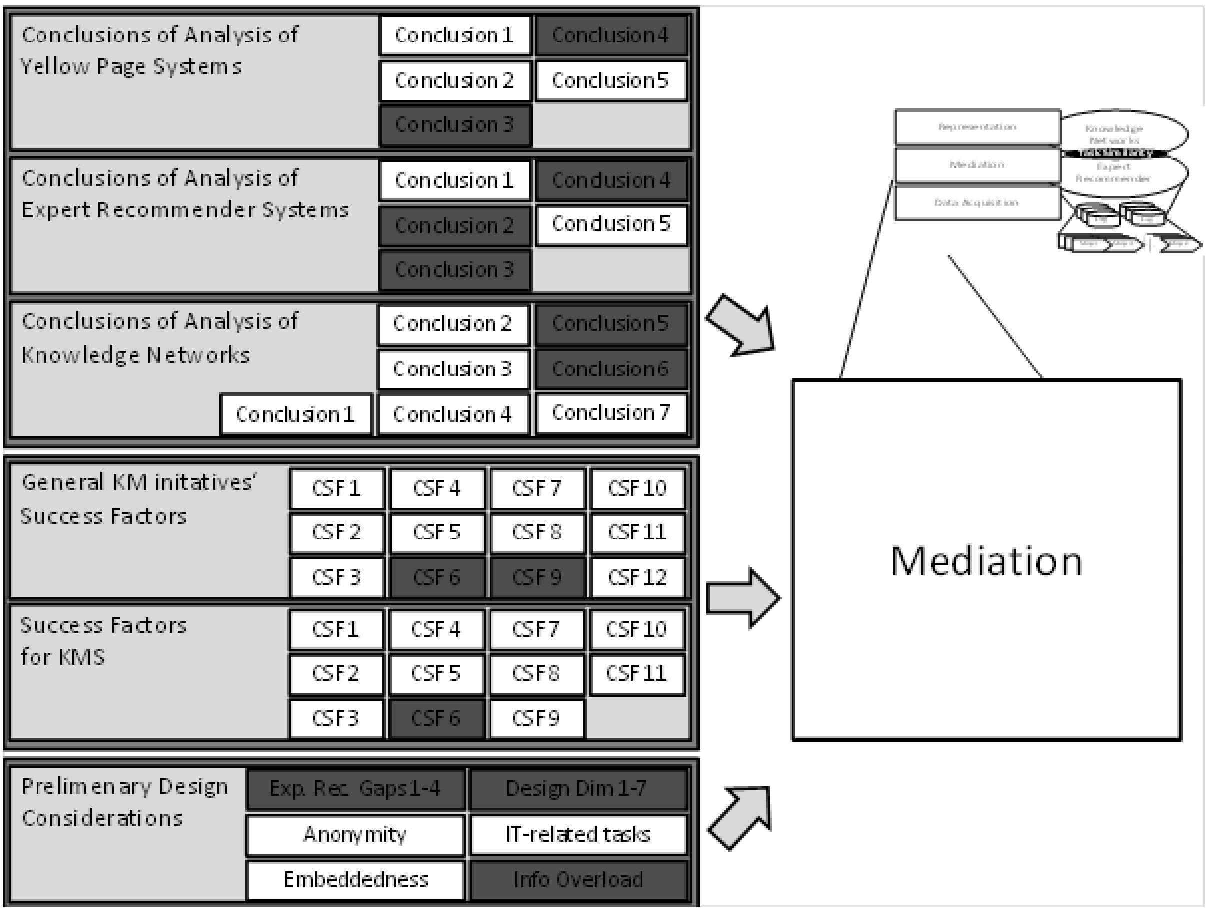


Figure 5-11 ***Influencing factors for the building block Mediation***
Source: Own illustration

5.3.3 Data Acquisition

Pursuing the goal of combining Knowledge Networks with ERS, similarity of task contexts has been proposed as suitable "glue" between those two concepts. To be able to use task contexts, appropriate data sources are necessary to determine the tasks that individuals executed and subsequently to determine their overlap. This is part of the lowest building block of this thesis approach as also depicted in Figure 5-9: *Data Acquisition*. The analysis of disadvantages of contemporary approaches for connecting individuals for knowledge exchange (see chapter 3) showed that effort for the individual needs to be as low as possible, for initial setup as well as for maintenance. This fact also reappeared as part of the general KM initiatives' CSF *motivation* and previously in this chapter again, where the reduction of *information overload* was discussed. Along with the observation that most tasks of a knowledge worker are supported by or result in an entry into IS, this thesis' approach focuses on analyzing tasks that are reflected in IS (see 5.1.1). More specifically, the results of these tasks that can be found in the log files of the respective IS will form the basis for the determination of similar tasks as indicated on the lower right part of Figure 5-9. Within this layer, a number of challenges and design considerations have to be taken into account. In the following, those are described by referring to the source of challenge in brackets after each statement. Additionally, Figure 5-12 visualizes which influencing factors determine the design of the building block *Data Acquisition*.

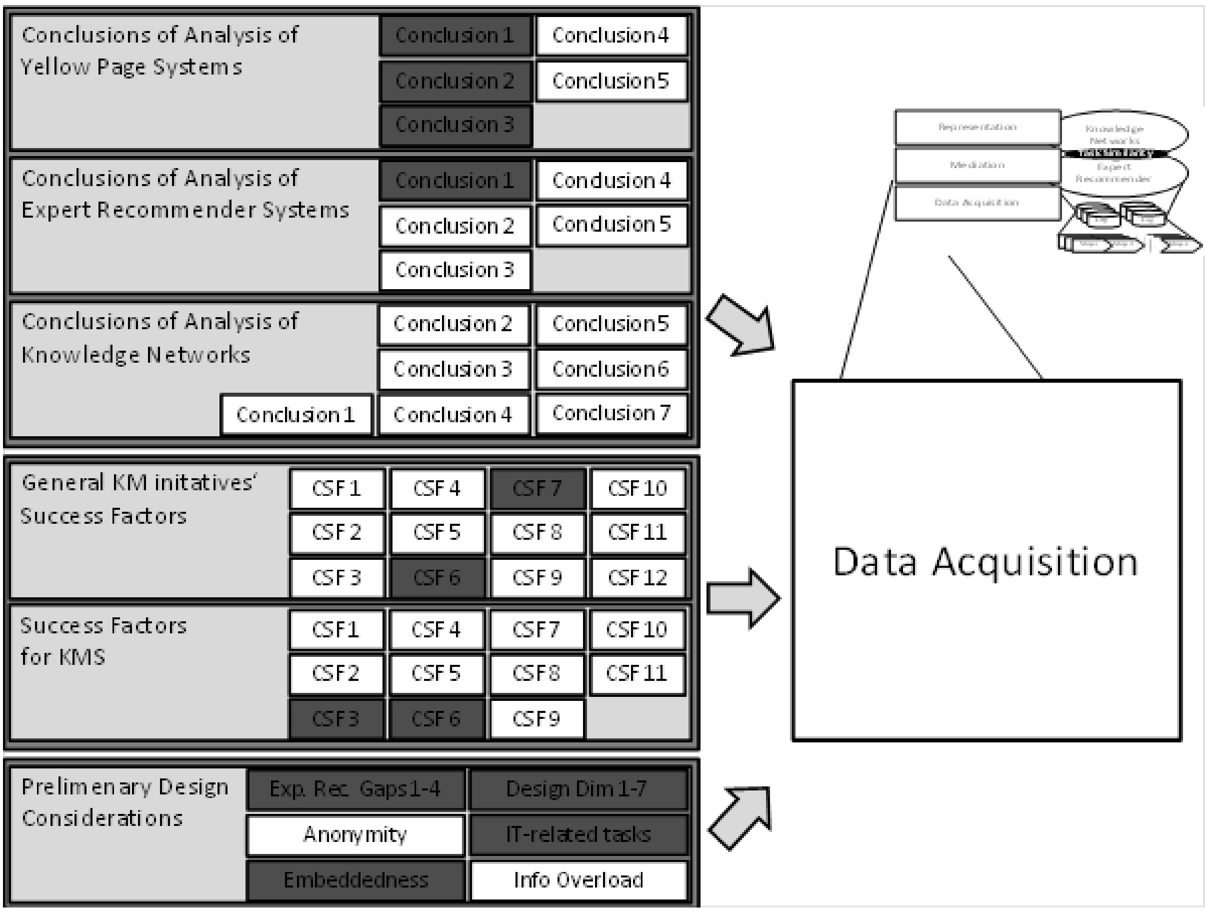


Figure 5-12 ***Influencing factors for the building block Data Acquisition***
Source: Own illustration

The data acquisition must be done in a way that is without any or with only **minimal burden** for the individual (Analysis of YPS: Conclusion 1). Further, a **suitable data entity** must be found that can be extracted and that reflects expertness and contextual overlap in a sufficiently exact way (Analysis of YPS: Conclusion 2 and Analysis of ERS: Conclusion 1 and Analysis of ERS: Conclusion 2 and ERS design dimension 1 and ERS design dimension 2). In addition, the data acquisition must be flexible enough to live with frequent changes of evidences of expertise and contextual overlap and consequently **frequent changes in the explicit or implicit model of users** (Analysis of YPS: Conclusion 3). The entities extracted during data acquisition must allow a later setup of expertise representations or **knowledge structures to which the users are already acquainted** (general KM CSF 6 and KMS CSF 6). Finally, the entities to extract should be **from operative IT systems** in which the individual users' requests will originate later, also allowing the aggregation of information from **more than one source** (general CSF 7 and KMS CSF 3 and ERS Gap 1 and ERS Gap 4 and preliminary design *Focus on IT Task* and preliminary design *Embeddedness*).

5.3.4 Remaining Influencing Factors

Having gathered all influencing factors for the design of this thesis' KM approach and taking the three conceptual layers for structuration, most factors could be matched and included. This will later inform the design of the mechanisms that support each layer as described in the following subchapters. As the investigation above showed, some influencing factors determine the behavior of more than one layer. For example, the requirement to take into account the context of a user whenever this user requests support is relevant for the *Mediation Layer* where the context is used to improve the matching of the individual with possible interaction partners. At the same time, the usage of the context is relevant in the *Representation Layer* where it may support the layout of the user interface in a context-dependant way or may automate or facilitate some interaction steps due to the knowledge about the current context.

On the other hand, when inspecting the influence of the identified factors on the approach's three layers, it turns out that some influencing factors have not been "used" in the design of this thesis' approach. For example, the general KM initiative CSF 2 – *Management Leadership and Support* – has not been addressed in any of the three layers. This influencing factor, as well as others like the alignment of *KM strategy* with the business strategy are rooted in the company-specific context and are outside the scope of the supporting mechanism itself that is designed. Nevertheless, they play a large role when the supporting mechanism is implemented in a concrete organizational setting. Those factors are therefore summarized into the encompassing *organizational context* as illustrated in Figure 5-13. The influencing factors in white boxes have been addressed by one of the layers, the light grey ones influence more than one layer of the approach. The dark grey ones are not addressed in the layers and hence are part of the encompassing organizational context that has to be adapted depending on the implementation scenario.

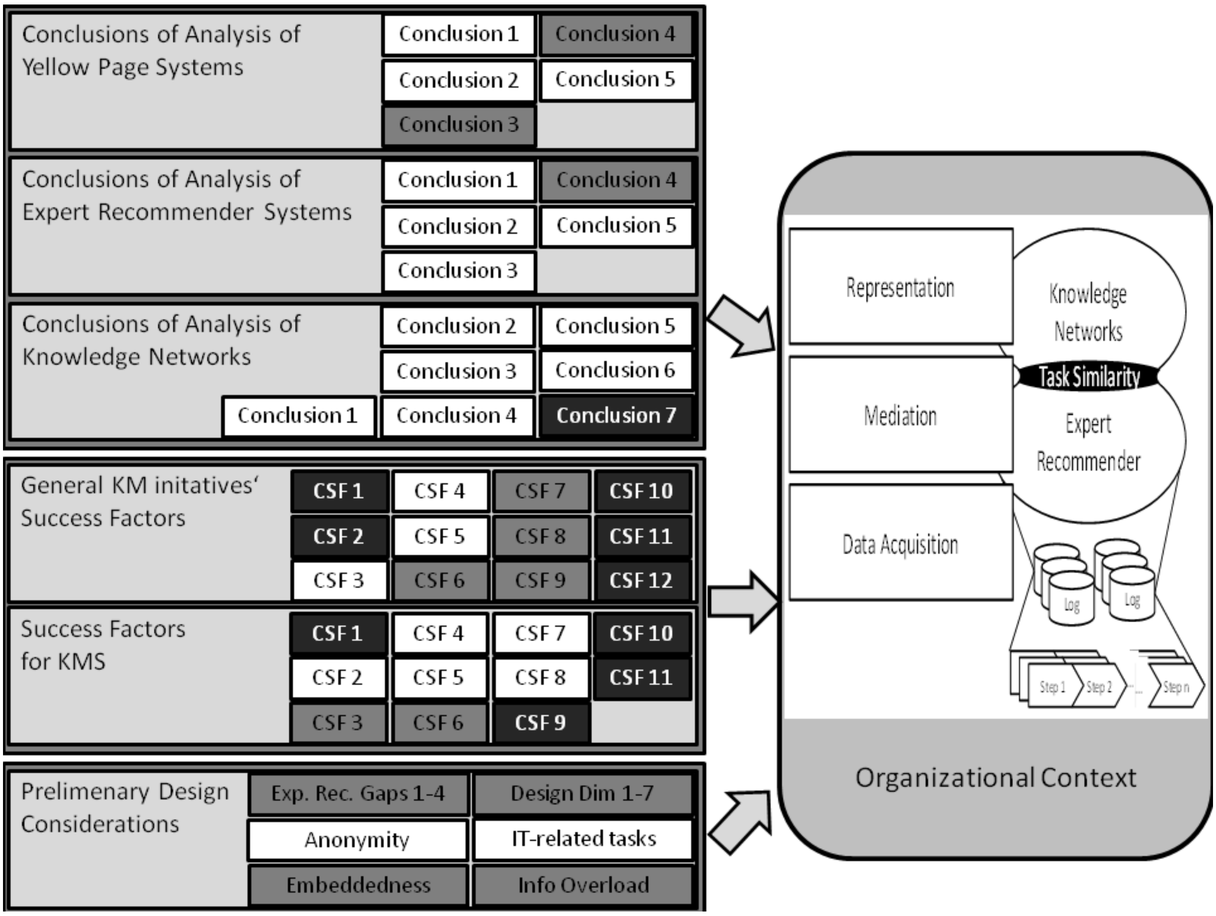


Figure 5-13 ***Remaining influencing factors and the encompassing organizational context***
Source: Own illustration

In the following, the *Organizational Context* as it relates to this thesis' approach is described. The way in which a company does business determines its strategy and consequently also the *KM strategy* that an initiative should follow. This thesis assumes a people-to-people KM strategy, i.e. knowledge is transferred through socialization processes, but does not make any more restrictions on how to leverage benefits of the interactions. Further restrictions are not possible as the relationship between KM and business itself cannot be generalized – there is not one truth but depending on the company the relationship may be diverse (general KM CSF 1). Further, while *Support from Management* for the adoption and prolonged use of this thesis' KM approach will be necessary, there is no way to support this from the mechanisms inherent to the approach. The support must be "granted" by the responsible managers (general KM CSF 2). Along with this, the *Resources* for the implementation and the maintenance of this thesis' approach in a concrete organization must be made available by the organization (general CSF 12). The approach may not change this fact but it includes features to decrease the effort of implementing and maintaining and hence reduces the amount of necessary resources (see chapter 6). Further, during the determination of CSF it turned out that successful KM initiatives integrate KM with *HCM* (general KM CSF 11). The details of this integration are part of a company's strategy and hence are part of the organizational context in the black-box view of the approach. This thesis' approach may however support some of the HCM tasks that relate to KM. For example, a tighter integration into social structures may be fostered by this thesis' approach and may in turn lead to higher retention rates. In addition, by better understanding which tasks are executed and discussed among the employees the necessity for professional development measures may become apparent. The final general KM success factor that has not found its way into the three layers is *Training* (general KM CSF 10). This factor was described as comprising different facets such as informing about organizational changes due to new roles and new incentive structures as well as informing about the KM initiative itself. A final aspect was the effect that trainings have on the affected individuals in the use of the new systems, hence in the implementation of this thesis approach.

Also, when considering the implementation process of this thesis' approach, there are some influencing factors that can be regarded as part of the *Organizational Context*. The selection of a proper *Software Platform* is a strategic decision to make. This thesis' approach offers a blueprint for this decision, but not a system itself (KMS CSF 1). Along this line, ensuring *Availability and Accessibility* certainly influences the adoption and perception of the KM approach described in this thesis, but depends on the hardware and software, such as servers, firewalls and backup systems that the implementing organization uses, but not on the general approach itself and hence is regarded as part of the organizational context (KMS CSF 9). The *Effort for Maintenance* also strongly depends on the concrete implementation, e.g. the way in which data is acquired (KMS CSF 10). Finally, the *Implementation Process* itself is a tactical or possibly even strategic decision that influences the user's acceptance and should be in line with the company culture and strategy. In general, this thesis' approach is agnostic of the implementation strategy. However, it may only provide benefit by fostering knowledge exchange if there is no initial aversion to it due to the implementation strategy and hence implementation processes that force the approach onto the users are less likely to yield benefit. Also, it should be clear that interpersonal exchange cannot be mandated and needs time to develop.

In the following, the discussion will focus on the three layers in turn, leaving the black-box view of this chapter, to understand how each layer's challenges are addressed and dealt with.

5.4 Representation Layer

The *Representation Layer* uses the context-dependent connections between individuals as determined by the *Mediation Layer* to actually route requests and answers between individuals and this way creates and intensifies a knowledge network between them. For an individual, the usage experience strongly resembles that of a knowledge network. In the black-box view on the *Representation Layer* in section 5.3.1 the following functionalities that the layer should support have been carved out:

- Integration into operative system where knowledge seeking originates
- Awareness of knowledge seeker's context
- Contribution of one-on-one interactions to more global learning
- Structured representation that is natural for users
- Management of content over time
- Finding already existing interaction possibilities
- Formation of smaller (sub-)networks
- Support for formation of new networks
- Increase of likelihood for contribution
- Enabling different levels of anonymity
- Support of different KM roles
- Support of different ways of incentifying
- Inclusion of a proper activity measurement

These functionalities concern different categories of abilities that are provided in the representation layer. Some refer to the *Management of the Network Size*, dealing with initialization and management of growth. Others concern the general *Management of the Network* and its supportive facilities for the management. The last group of requirements deals with the *User Interaction* and the representation of the network towards the user. In the following, each of these categories of abilities will be discussed in turn.

5.4.1 Network Size Management

Within the representation layer’s category *Network Size Management* the following functionalities are subsumed as they all relate to the size of a network:

- Finding already existing interaction possibilities
- Formation of smaller (sub-)networks
- Support of formation of new networks

The more established knowledge networks[5] become within an organization the more one or both of the following phenomena will incur: The size of a single network may increase to a level that is above what is manageable in terms of relations to other individuals and knowing

[5] In the following when refering to knowledge networks, (classical) communities of practice are excluded, hence only distributed networks of practice or electronic networks of practice are subsumed by this term

about the others. In addition, there may be more and more networks spurring in the company that focus around different topics. Here again, it may happen that the number of networks exceeds a level that is manageable for an individual with limited amount of time to spare. Taking again the example of section 3.4.7.3, an electronic network of practice operated by SAP, one can observe both problems. While the number of members is unknown, one forum that the author is member of has, at the date of writing, more than 25.000 threads with together more than 130.000 single messages and on average 28 new messages a day are posted. Clearly more than one can handle – both in terms of knowledge that is hidden somewhere in previous threads as well as in new messages per day. Also the number of 230 networks that the exemplary SAP network features is more than one can keep track of. When someone has a problem and wants to post a message, a first hurdle therefore may be to find the right place to request help. The effort to find a suitable "outlet" may be a hindrance to ask for help at all. Also it is a hindrance for someone who is in principle willing to help but not willing to take the effort of "searching" for requests that he could help with. Finally, it is a hindrance for peripheral members – lurkers that read messages, increase their expertise and this way increase the company's overall ability to act. As the large number of possibilities increases their effort for browsing forums that could matter to them, it consequently decreases the effects of learning effects for the peripheral members.

Consequently, this thesis' approach should limit this problem. There are different ways in which it can help. Firstly, the effort for finding the right place on the requestor's side is removed. In the *Representation Layer*, a user enters a request and, taking his current and previous task executions into account, suitable interaction partners are determined automatically. This is a functionality provided by the *Mediation Layer* that is visible to the benefiting individual through the *Representation Layer*. Secondly, possible interaction partners – experts for this interaction – are informed automatically and anonymously taking into account their task context as well. This way, only requests that are potentially relevant with respect to the work of the request-receiving individual are sent. If the receiver does not wish or cannot respond adequately to the request he may still "lurk" in the discussion to learn. As his context was taken as criterion for selection, there is a high chance that the lurker is interested in the discussion. Furthermore, taking the empirical results of section 5.2.1, this thesis' approach may increase social capital and hence positively influence contribution behavior. Law and Chang (2008) could show that more central users contribute more knowledge. This thesis' approach deliberately generates opportunities for individuals to become more central by responding to requests, that deal with tasks they perform and hence tasks that they should be naturally more interested.

There are other ways in which this thesis' approach provides support for knowledge exchange: In situations where knowledge networks start to lose value due to too large size and hence number of interactions (*information overload*) or too diverse topics and diverse members' interest (loosing focus and hence ability to identify with the network). A possibility is to form (sub-)networks, that has been shown to foster social capital (see Law, Chang 2008 and the discussion in section 5.2.1). The empirical evidence showed that more focused groups foster identification with the (sub-)network and the language is more focused to what is common for its members. Both these effects increase social capital and hence, relying on SCT, result in more knowledge contribution. In addition, more focused (sub-)networks have less interactions and those that do happen are more relevant to its members and hence (unnecessary) information overload is less.

While Law and Chang (2008) described and verified these properties in their empirical study, they did not indicate how sub-community formation can be done. Especially in larger networks, it may not be obvious which sub-networks, given only the initial large network to split, would be a good selection. In this thesis' approach, requests are created with relation to a certain task context, which is also associated to the discussion that results from the request. A core service provided by the *Mediation Layer* is to determine task context overlap. This is also possible for the task contexts that are assigned to the requests that initiated a discussions. When forming sub-networks, this thesis' approach offers the ability to indicate minimum levels of context overlap, i.e. minimum values of similarity of the task contexts that are associated with the resulting discussions. This way the number of messages is limited in a way that those discussions remain together, that have high contextual overlap. As a result identification with the overall network may be increased as the included sub-networks become more focused to what the individuals do. Schematically this is depicted in Figure 5-14. Similarly, the same mechanism may be used to form new networks, when applied in the inverse. There may be a situation in which predominantly unrelated discussions are the primary means of interaction, e.g. because an organization used an approach more similar to an ERS with some structuring functionality attached. If this organization now wishes to foster the formation of networks, finding those individuals that have requested help for and provided help on similar tasks is a suitable initial nucleus for new communities to be formed. This allows to identify suitable candidates and to forgo creating communities that are "[..] created in a vacuum." (Wenger 2000, p. 144). Crystallizing networks around tasks also increases the chance that discussions are more specific and more concrete with respect to the task. According to the theoretically derived model of Olivera et al. (2008) more concrete requests will increase motivation to respond to someone's request and hence knowledge transfer (see section 5.2.2, hypotheses 4,5,11,13).

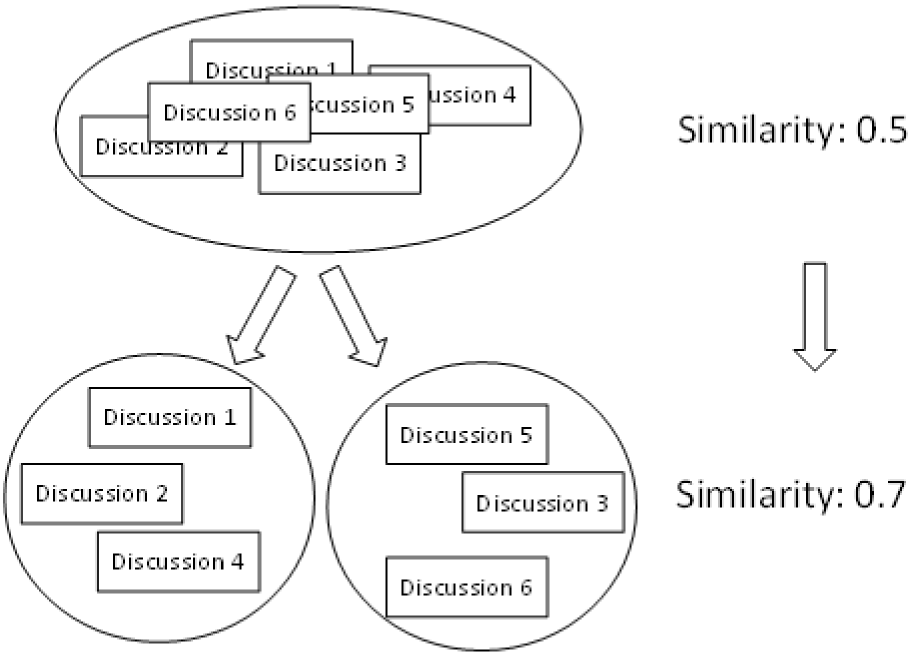


Figure 5-14 ***Sub-network formation using minimal similarity values***
Source: Own illustration

Finally, the meta-information about the context from which a discussion started may be used as supporting information for an individual's search for knowledge stored in the network's interactions. According to the theoretically derived model of Olivera et al. (2008) helpful search facilities will increase motivation to respond to someone's request and hence knowledge transfer (see section 5.2.2, hypothesis 14). A contextualized search like the one made possible with this thesis approach represents a helpful search facility.

5.4.2 Manage the Network

Within the representation layer's category *Manage the Network* the following functionalities are subsumed as they all relate to the "provider", i.e. that are the organization's interests in managing the approach and more specifically the *Representation Layer* thereof:

- Contribution of one-on-one interactions to more global learning
- Support of different KM roles
- Enabling different levels of anonymity
- Support of different ways of incentifying
- Inclusion of a proper activity measurement

By integrating expert recommendation facilities that route requests to suitable interaction partners, this thesis' approach must cope with the natural challenge that one-on-one interactions are created that others may not benefit from. From an organizational point of view this has the disadvantage of leading to an unmanaged network of knowledge flows (Oldigs-Kerber

2007, p. 67) and organizational learning is restricted to dyadic benefits instead of multiple possibilities to spread knowledge. This thesis' approach, therefore, uses two measures to route requests to experts while still enabling others to benefit from the resulting interactions. First, the determination of suitable interactions partners does not yield only one (presumably best fitting) interaction partner to whom the requests of a knowledge seeker is sent. Instead, a configurable number of possible interaction partners is compiled that are all anonymously put on a virtual list of interaction partners and informed about the request, similar to the *dynamic mailing list* approach described in (Ye et al. 2008). Taking into account that the determination of this virtual list of interaction partners is formed also by considering task context overlap, the other interaction partners are likely to benefit also from just lurking, i.e. "listening" to the resulting conversation as it affects their previous or current tasks. Second, now in correspondence to discussion forums of knowledge networks, the representation layer features an additional functionality: Interactions are stored in a forum-like structure for later browsing. To remedy the problems of *information overload* that might occur when many interactions are stored this way, the presentation of the stored interactions should be context-specific and in a "language" that is typical for the individuals using the approach. Hence, the interactions will be stored according to task or, if available, a taxonomy of tasks, e.g. task group, tasks and subtasks.

Being a hybrid approach and hence partly a knowledge network, this thesis' approach must also take into account that the social fabric that will evolve follows the rules of virtual communities, as knowledge networks are a subset thereof. In virtual communities, not every member is equally active. For example, Wenger et al. (2002, p. 56) note that only 10-15 % of a community's members are core members and hence very active. Nevertheless, those members carry the success of the community. Another important kind of community member are coordinators who help the community to focus, maintain relationships and develop its practice (Wenger et al. 2002, pp. 71,73). In this thesis' approach, their most important task will be to tackle the initiation challenge, i.e. to provide a critical mass of useful knowledge resources of sufficient quality during the launch of the KM initiative to not fall victim to the death spiral of KM initiatives (Probst et al. 2006). In the approach at hand, this translates to making sure that the first requests are responded to quickly to increase trust in the approach. In addition, coordinators are the main contact persons for any questions to the approach and especially to communicate and possibly enforce norms of cooperation that evolved or that have been determined by the organization. The influence of norms of cooperation on social capital and hence contribution behavior has been argued for in section 4.2.11.

Earlier during the discussion of the design considerations for this thesis' approach, it surfaced that for the end user the result may resemble a distributed network of practice – at least some part of it – or electronic networks of practice. Besides other differences (see sections

3.4.4 and 3.4.5 for details), the degree of anonymity is different in both. Therefore, instantiations of this thesis' approach should either support anonymity by design, or should offer the possibility to show (real) names of the individuals.

While earlier, the thesis argued that the benefit that a KM approach can provide for an individual should be the incentive itself in the long run (see section 4.2.11), especially to jump-start new initiatives other forms of motivational aids are helpful. The theoretical and empirical bases analyzed earlier in this thesis determined that *reputation* gained through contributions is especially important (see section 5.2.1). If the approach is used to resemble a distributed network of practice and hence members are known to each other and their identity is expressed, reputation is naturally increased when individuals engage in pro-social and intelligent interactions. If the approach is used to resemble electronic networks of practice where individuals are typically anonymous and unknown to each other, this may be more complicated. On the one hand, users may deliberately include their names and affiliations in the request itself if they like. From a technical point of view, this can be supported by enabling the users to choose whether they want their requests to be seen under a pseudonym or with their real name. On the other hand, to balance between anonymity and gaining reputation the approach may allow to have anonymous users, but their achievements, i.e. messages posted and positive feedback they received, can be displayed disentangled from the actual contributions – the interaction itself is still anonymous.

Finally, for managing the network, activity measures need to be provided that inform the relevant stakeholder, e.g. the coordinators, about the liveliness of the network. In addition, information about which kinds of tasks are discussed most often and hence which knowledge is (currently) important to the individuals may be determined this way (Davenport 1997, p. 69). Different measures have been proposed for this purpose ranging from quantitative numbers of exchanged messages or requests to numbers of appreciations of other's contributions that reflect quality. Other measures from social network analysis like centrality, betweenness or connectedness were suggested also as being suitable proxies for the liveliness and quality of a knowledge network. It is outside the scope of this thesis to evaluate which best reflects "real activity" in the proposed approach and the definition of what is success, also of knowledge networks or ERS, is still an open research question (see discussion in section 4.2.7).

5.4.3 User Interaction

Within the representation layer's category *User Interaction* the following functionalities are subsumed as they all relate to the "interface" towards the individuals using the approach:

- Awareness of knowledge seeker's context
- Integration into operative system where knowledge seeking originates

- Structured representation that is natural for users
- Management of content over time
- Increase of likelihood for contribution

The knowledge seeker's current task context should be taken into account for various purposes. First, the context can be used to help the knowledge seeker find relevant knowledge that helps with a current problem and that might be contained in the recordings of previous interactions (see also discussion in section 5.4.1). Second, the task context can be used to convey extra (meta-)data that the *Mediation Layer* may use to find suitable interaction partners.

In the previous discussion about CSFs for KMS it surfaced that tight integration with existing business processes contributes to acceptance (see section 4.3.4). In section 5.1.5 this aspect was described in greater detail, highlighting that integration should not only focus on the business process but also on the system that knowledge workers use for the execution of their tasks. This asks for two related things. First, it should ideally be possible to seek necessary knowledge *directly from the operational system*, which means that operational systems need to be adapted or at least a linking mechanism needs to be integrated into the system. In a less optimal situation, the system cannot be adapted and the act of seeking knowledge is performed in a different system. In this case, it is however still necessary to extract the current task context of the user which can typically be done by accessing log files of operative systems, hence without adaptation (see chapter 5.6 for a discussion of data extraction challenges). Second, integration into operational systems also encompasses the *notification of possible responders*, again ideally within the operational system. If a user is currently working at the same or a similar task as a knowledge seeker and is notified about a request immediately while in the same task context, the perceived social presence is especially high. According to the theoretically derived model of Olivera et al. (2008), this increases the motivation to contribute knowledge (see section 5.2.2).

This thesis' approach supports structuring the representation in the user's *natural manner* in different ways. First, knowledge seeker's requests are visually amended with the current task to reflect context using the task names or descriptions common to the individual. This way, both the knowledge seeker but especially the possible respondents have an additional stimulus that allows them to quickly determine first approximations of whether or not there will be a high overlap between necessary and available knowledge to support the knowledge seeker. This step is performed by every possible respondent according to *Problem Solving Theory* (Newell et al. 1972) and its implications to the model of Olivera et al. (2008). Hence, if this thesis' approach reduces this effort, the collective attention that is consumed is reduced and the overall interaction is more efficient. Second, and in line with the discussion in section 5.4.1 about reuse potentials, previous interactions are sorted and categorized according to task or task hierarchies that are represented in the natural “language” that is typical for the individuals using the approach.

The elicitation of CSFs hinted at the relevance of offering the ability to manage content in terms of rating. Here, the users should be enabled to indicate what is relevant for them and what is no longer. In the context of this thesis’ approach, this is naturally no challenge for the requestor. The current task context is extracted from the operational systems that is used and hence the meta-information about the knowledge seeker’s context is automatically up-to-date. The user can utilize this meta-information to determine the relevance of interactions in a first quick estimation. However, individuals in the role of knowledge bearers might need adaptability in terms of content that they are interested in. For example, someone might not be interested any longer in stepping into discussion concerning a certain task. *Aging*, i.e. gradually decreasing importance of single tasks over time, within the collection of the knowledge bearer’s task context naturally reduces the chance of receiving requests with respect to tasks that the knowledge bearer no longer performs. However, it may be possible that someone is still performing a task but does not want to receive requests for support anyway. In this case a mechanism indicating that no more requests for this topic, i.e. task should be send is a suitable measure. Likewise a knowledge bearer might not want to interact with a certain individual in which case a *negative list* for this user is a viable mechanism.

Finally, the user interaction component of the *Representation Layer* needs to incorporate means to *increase likelihood of contribution*. This can be achieved in different ways. First, as knowledge seekers can concentrate on formulating a request for knowledge, they are relieved of the burden of finding suitable interaction partners. While this inflicts less effort for the knowledge seeker, there is also another benefit. In an empirical study Vitari and Ravarini (2009, p. 9) found that there is “ [..] a barrier to asking for the provision of an ERS [expert recommendation service; remark of the author], and more generally, to asking for help from members about whom the seeker has little or no knowledge.”. This means without the concept of relieving the knowledge seeker of the burden, not only would the effort be higher but potentially the request would not be put forth at all. This thesis’ approach increases the likelihood of contribution also on the other side – the responder’s side. Because for the matching, contextual overlap is utilized but also levels of expertise influence the selection, the potential responders are asked about topics that they are likely able to help with and also that they are likely to care about. In accordance with the theoretically derived model of Olivera et al. (2008) this increases the likelihood of contribution (see section 5.4.3, hypothesis 3). Additionally, a mechanism for appreciation of other user’s contribution in a discussion is part of the approach. This increases a knowledge bearer’s ability to build up *reputation*, which was shown to be a strong motivation factor (see section 5.2.1).

5.4.4 Summary

Summarizing the discussion about the *Representation Layer*, one can conclude that there are three categories of functionalities that should be supported: *Network Size Management*, *General Management* aspects of the network and *User Interaction*. Network size can be managed using task context information for crystallizing new networks around common context, separating too large networks into meaningful smaller networks or by finding suitable networks contingent on task context. The management of the network is supported in various ways. A mechanism for spreading dyadic learning effects to a larger group of stakeholders has been discussed that capitalized on the inclusion of possibly interested individuals and storage of previous interactions. Furthermore, the roles were described that individuals may assume in the networks forming upon the use of this thesis' approach. They included knowledge seeker, knowledge bearer and coordinator, the later caring for the liveliness of the network especially in its early stage. In addition, the discussion included the effects of anonymity on the network's features and how the approach addresses this fact. Anonymity can be granted, the choice can be given to the user to be anonymous or not and also the reputation effect and the actual interactions can be treated with different levels of anonymity. Along this line, motivation effects have been discussed among which reputation play a significant role. A discussion about possible operationalizations of activity measurement was a next part of the chapter, concluding that there are no dependable results as to which measure is best suitable. The interaction with individuals mediated by this thesis' approach was inspected next. The incorporation of the user context is taken into consideration as is the question how interactions with operational IS can be tailored. Then, the question how to structure interactions within the network is discussed pointing towards the adoption of established terminology of the source systems, that can be used for guiding personalized content views.

5.5 Mediation Layer

In the *Mediation Layer*, context and expertise models of the individuals are used to create context-dependent connections between them when an individual seeks knowledge exchange with other individuals. In the black-box view on the *Mediation Layer* in section 5.3.2 the following functionalities that the layer should support have been carved out:

- Operation on explicit or implicit models of user context and expertise
- Matching of individuals based on expertise but also contextual overlap
- Taking into account the context of the knowledge-seeking individual
- Minimal or no effort for the user
- Attention to limiting the consumed collective attention to a necessary minimum

In the following, the question how connections between knowledge seeker and knowledge bearer may be obtained will guide the discussion. First, a conceptual differentiation between expertise and context model will take center stage. Next, the term *similarity* will be discussed in depth along with possible concepts that may be used to determine similarity in this thesis' context. Finally, the resulting effects of this approach's usage on the *collective attention economy* complement the discussion on how to mediate individuals for knowledge transfer.

5.5.1 Expertise vs. Context Overlap

When an individual is confronted with a challenge that renders the support of another individual advantageous, because the seeking individual does not possess the necessary expertise, or if the sought for individual may deliver support with less cost for the seeking individual (e.g. no searching for a solution necessary), the question arises who to contact for help. Typical ERS (see section 3.3.2) try to support the seeking individual by providing "know-who" capabilities, i.e. they indicate who knows much about a certain topic. They then return a sometimes ranked list that contains the most knowledgeable individuals for a typically textual query. However, the most knowledgeable interaction partner may not be the "best" interaction partner due to two reasons: inefficient interactions and less motivation to interact (see section 3.3.4).

Taking into account the data that will be gathered in the *Data Acquisition Layer* – tasks and their contextual information like time and executor – one can create an implicit model of a user. Taking a closer look at the data, the implicit model that this thesis' approach capitalizes on, can be used to extract two different kinds of information, namely *expertise* and *context* of the individual. In the following, both aspects will be described in elaborated detail by referring to socio-psychological research results and incorporating those results into this thesis' approach.

The approach at hand, tries to connect a knowledge seeker having a requests for help with some knowledge bearers, where typically the knowledge seeker has less expertise in the relevant topic than the "expert". This is similar to expert-layperson interactions, where also one individual has presumably higher expertise in an area than the other. When an expert interacts with a layperson, the communication between the two is typically not easy as they have different mental structures, i.e. different *mental models* (Bromme et al. 2004). To store larger expertise and knowledge about a topic, individuals "compress" their knowledge about a topic by aggregating single facts into larger chunks and then using abstractions to condense their knowledge. When a layperson asks an expert for help, the expert is faced with the challenge to undo his learning, i.e. to unpack the knowledge he has into more "profane" levels of abstraction so that he reaches a level of understanding that is adequate for interacting with the layperson. This is a non-trivial task for the expert and requires high mental effort. However, this is necessary to reach a *common ground for communication* (Clark 2007) – an overlap of the two individual's subjective mental models that allows to interact with a sufficiently high probability of understanding what the other individual says. Hence, asking the most knowledgeable person on a certain topic may induce very high costs, also on the side of the expert, which is problematic as it is typically those experts that are short in time. Additionally, Bromme et al. (1996) note that an expert's mental representation of knowledge about a topic is *situated* and hence structured to the expert's *specific work context*. Consequently, if the work context of two individuals has a larger overlap, the mental structures about a topic will likely be similar and communication about the topic is easier for the two individuals. This is in line with Ackerman's analysis of how networks of individuals operate. He states that "[...] the people nearby the user are the most likely to understand the user's context. Since they know the user, they can also make best judgments about how to present the answer." (Ackerman, Halverson 2004, p. 9). However, if the individuals around the knowledge-seeker cannot help, picking up on the discussion from before, the next best equivalent to “people nearby” is to ask those with still a high contextual overlap, i.e. that have the same current or previous work context. Taken together with the argument before, that the effort for an expert to find a common ground for interaction is demanding, one can conclude that similarity in work may be a good means to connect individuals for fruitful interaction. The similarity in work translates to similarity of executed tasks in this thesis' approach.

In addition to finding interaction partners that are capable of communicating with the knowledge seeking individual, another prerequisite is to find interaction partners that likely *want* to interact with a knowledge seeker. In socio-psychology, the notion of *interpersonal attraction* has been studied intensively trying to answer the question “who do we like”. Having a positive attitude towards an interaction partner, i.e. “liking” the other one, strongly increases the motivation to engage in mutual interaction and knowledge exchange. Zimbardo (1983, p. 606) notes that one of the most clear results from socio-psychology is that one likes others that are similar to oneself. Further support for this can be found in a ground-laying experiment by Byrne (1971) showing that similar attitudes are a predictor of interpersonal attraction. Zimbardo (1983, p. 607) notes that using the *Social Exchange Theory* (Homans 1958) is one way to describe why attraction to similar individuals is higher. It postulates that individuals try to increase benefit and limit costs in social interactions of any kind. Zimbardo reasons that when two individuals share aspects then the costs in terms of time and effort are lower and the benefits are higher. Picking another detail of this field of research, Ziegler and Golbeck (2007) explored the relation between *similarity of interest* and *trust* between individuals in online recommender communities and could show that those two concepts are correlated when looking at a specific domain – recommendation communities in their case. Taking these results into account suggests that individuals with similar work context are more likely to interact because they are bound by commonalities – their similar work – and hence are likely to have more positive expectations towards the interaction partner. Again, in this thesis' context, similar work translates to similar tasks.

The previous paragraphs looked at theories that underline why finding the most knowledgeable individual is not necessarily the best choice for a knowledge seeker but instead a contextual overlap can increase the benefit. On the other hand, one does not ask others merely because one likes the other individual or because one can talk in the "same language" to this individual. Still the goal is to find individuals that exhibit suitable expertise in the area that the knowledge-seeking individual needs support in. As Terveen stated with respect to recommendations: "While getting recommendations from somebody with similar tastes is a good start, you might also want something else: that the person making the recommendation is an expert on these topics." (Terveen, Hill 2002). While similar taste is not the same as similar work and tasks, the argument is still transferable to this thesis' context: the individual that is requested for help should be expert enough to be able to help.

Acknowledging the importance of both aspects – level of expertise and contextual overlap – there are different ways of combining them. Figure 5-15 depicts the principle ways of distinguishing prospective interaction partners taking *expertise and contextual overlap* into account. As a first way to address the challenge, one may use a minimum value of expertness for any individual that should be connected in response to a knowledge-seeking individual's request. This minimum value can be "absolute", e.g. only individuals who have performed a task ten or more times, or it can be relative to the knowledge seeker, e.g. only individuals that have performed a task at least five times more than the knowledge seeker has. If an individual is below this threshold (rectangle II and IV in Figure 5-15) one would not try to connect this individual with the knowledge seeker. The set of individuals that are above the threshold, may all be connected or the second dimension, the *contextual overlap*, can be used to rank the individuals and only try to connect some or maybe only one with the knowledge seeker according to this ranking. Similarly, in a second strategy, one may use a minimum value of contextual overlap to find those individuals that are possibly willing to interact with the knowledge seeker and that also can do so with likely little effort because their mental models are similar. Naturally, the minimum value is always relative to the knowledge-seeker's context. Again, if an individual is below the threshold (rectangle III and IV in Figure 5-15) one would not try to connect this individual with the knowledge seeker. As before, the set of individuals above the threshold may all be connected or the first dimension, the *level of expertise*, may be used to rank the individuals and connect only a subset. In a third strategy, one may use both thresholds so that only those individuals will be connected to the knowledge seeker, that are sufficiently knowledgeable and that are sufficiently alike to the knowledge seeker (rectangle I in Figure 5-15) while all other individuals are not taken into account (rectangle II, III and IV in

Figure 5-15). In a fourth strategy, one may use a combined threshold. Individuals that are very similar to the knowledge seeker but have only low levels of expertise may be suitable, as in the opposite case, individuals that do not share much commonalities with the knowledge seeker but are very knowledgeable may be suitable candidates as well. Therefore, in both cases one dimension might compensate the lack in the other. However, individuals that neither have sufficient contextual overlap nor sufficient levels of expertise (darker area in the lower left part of Figure 5-15) are those that should not be connected to the knowledge seeker in this strategy.

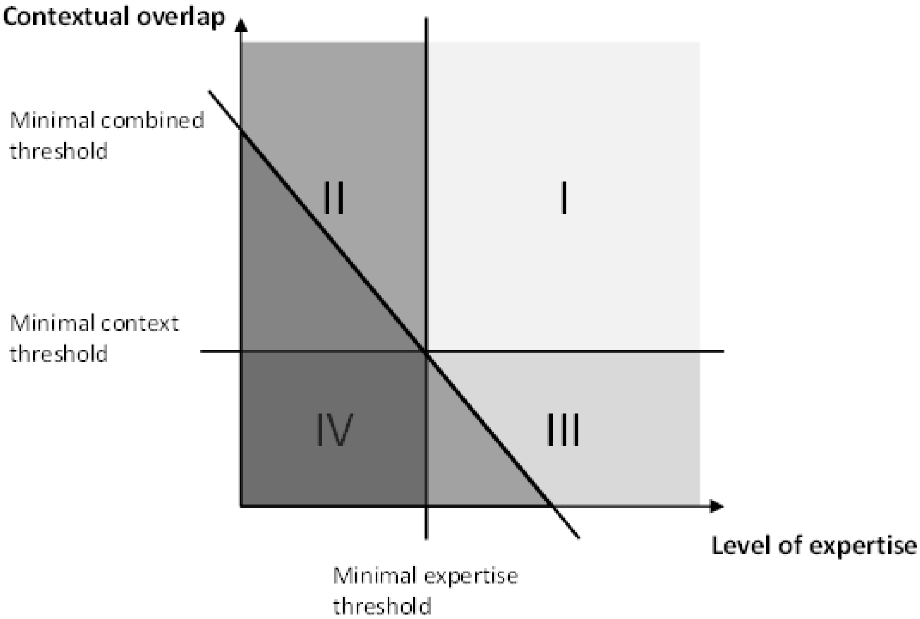


Figure 5-15 ***Contextual overlap and level of expertise matrix***
Source: Own illustration

The use of threshold values and the choice between the four strategies that were described is a fundamental part of how to determine suitable interaction partners. However, as Seid and Kobsa (2003, pp. 10–12) noted in their description of systems that work with recommendations, a suitable approach must consider how to include *adaptability* for their users to learn from explicit or implicit feedback. Consequently, in addition to the two fundamental dimensions that were described, the algorithm for matching individuals may include other measures to increase or decrease the set of possible interaction partners. For example, the matching algorithm may include negative and positive listings for filtering the result set. Individuals may want to indicate that they want to interact more or not any longer with a certain individual and put them on their positive or negative list. The matching algorithm can adapt the measure of suitability according to this extra information. Likewise, an implementation of the approach may offer the user the ability to indicate that he wishes to interact more with respect to a certain topic or may wish to not start interactions on this topic any longer.

5.5.2 Similarity as Mediator Between Knowledge Seeker and Expert

The previous section highlighted that contextual overlap is a crucial aspect of finding suitable interaction partners in this thesis' approach. It further argued that contextual overlap translates to having performed similar tasks in the (recent) past. However, just looking at the one task at hand may not contain sufficient information to state that someone is working in a similar way. In some situations it is rather the sets of tasks that better resembles the notion of similar work. Then again, it may be the case that it is not only what someone does that determines contextual overlap but also how, i.e. in which sequence. Obviously the term *similar* is used with a broad meaning in everyday language, but needs to be investigated more closely to make it operationalizable in this thesis' context and to use it for determining contextual overlap.

This section first provides definitions of the necessary terminology. The terms *similarity measure* and *distance measure* are defined, different kinds of measures are explained before similarity by use of feature vector and by distance determination are distinguished and levels of abstraction for similarity measures are discussed.

5.5.2.1 Definitions

In this section the most important similarity-related definitions are provided. The definitions of *similarity measure* and *distance measure* are given. Then different kinds of measures are distinguished, namely *kernels* and *metrics*. Then the two principal approaches for defining similarity, the transformation into a *vector space* and the direct calculation of distance are described. Similarity and distance are related measures that often determine the inverse of the other. They can be determined with a function that assess closeness - a kernel - or by a function that assesses distance - a metric. A metric has more stringent demands than a kernel, therefore a metric can always serve as basis for a kernel, the other way around is also possible under certain circumstances. Feature vectors are created by utilizing the most discriminating features of an object and projecting those into a vector space. Norms or angles can be used to assess similarity of vectors. Distance-based similarity is highly application-specific, often computationally costly but well-suited for its intent.

5.5.2.1.1 Similarity Measure

The task of a *similarity measure* is most obviously to measure the similarity between two elements. The input elements are of a certain *domain*, such as the domain of vectors or the domain of graphs. Thus, a similarity measure has, as a first property, a specific domain on which it can operate. A second property of a similarity measure is, that the value of the measure needs to be higher, the more similar two elements are. It is further desirable, yet not necessary, to require as a third property for the similarity measure to be normalized, thus, whenever two elements are equal the measure should have a value of 1. The final property of a similarity measure demands *symmetry*. Thus x_0 is as similar to x_1 as x_1 is similar to x_0.

5.5.2.1.2 Distance Measure

Most obviously a distance measure's task is to determine the distance between two elements. Therefore it is sometimes also called *discrimination function*, as it can be used to tell apart two elements. The distance of two elements corresponds to their similarity, in such a way that distance can be interpreted as "inverse-similarity". In other words, the more distant two elements are, the less similar they are and the other way around. There is no upper limit for a distance, yet there is a lower limit. The distance between two elements must be *equal or greater than zero*. If requiring a similarity measure's and a distance measure's maximum value to be 1, the correspondence between distance and similarity is as follows

$$\boldsymbol{value_{sim} = 1 - value_{distance}}$$

Analogously to the similarity measure, a distance measure needs to be *symmetric*.

5.5.2.2 Kinds of Similarity Measures

As stated above, similarity and distance are closely related concepts. Therefore, the methods of measuring either one of the two are also quite related and in many fields of application it is possible and feasible to choose either one of the two as basis. Depending on if one wants to measure the relatedness or the unrelatedness of two elements, one can either use a *kernel* for the first case or a *metric* (sometimes only a distance measure as defined below) in the second case. Note that it is also possible to use a (normalized) distance measure for measuring similarity and vice versa using the equality given above.

5.5.2.2.1 Kernel

A *kernel* is a function that takes two input elements, returning their relatedness. The minimum requirements for a kernel are as follows:

- $K(x_0, x_1) \geq 0$ (*positive semi-definite*)
- $K(x_0, x_1) = K(x_1, x_0)$ (*symmetry*)

where x_0 and x_1 are elements of the input domain.

Therefore, a kernel fulfills the requirements of a similarity measure, if it additionally returns higher values for more similar elements.

5.5.2.2.2 Metric

A *metric* is a function that takes two input elements, returning their distance or unrelatedness. The requirements for a metric are as follows

- $d(x_0, x_0) = 0$ (identical points have distance 0)
- $d(x_0, x_1) \geq 0$ (*non-negativity*)
- $d(x_0, x_1) = 0 \Leftarrow: x_0 = x_1$ (*definite*)
- $d(x_0, x_1) = d(x_1, x_0)$ (*symmetry*)
- $d(x_0, x_1) \leq d(x_0, x_2) + d(x_2, x_1)$ (*triangular inequality*)

where again x_0, x_1 and x_2 are elements from the input domain. If only the first two properties are satisfied the function is called a *distance measure*. If the function is not definite, but only semi-definite the functions is said to be a *semi-metric*. Thus, a metric is more specific and has more stringent requirements than a kernel. Furthermore, a metric is used to determine distance instead of similarity, thus the correspondence of similarity and distance has to be used to derive a similarity measure, as desired.

5.5.2.2.3 Relationship of Kernel Functions and Metric Functions

In general, kernels measure similarity while metrics measure distance. However, as stated above there is a close relationship between kernel functions and metric functions. It is always possible to derive a kernel function from a metric using e.g. the following equation.

$$\boldsymbol{K(x_0, x_1) = e^{-d(x_0, x_1)^2}}$$

If $K(x_0, x_0) \geq K(x_0, x_1)\ \forall\, x_0, x_1 \in inputDomain$, then it is also possible to derive a metric from a kernel using e.g. the following equation

$$\boldsymbol{d(x_0, x_1) = \sqrt{K(x_0, x_0) + K(x_1, x_1) - 2K(x_0, x_1)}}$$

5.5.2.3 Feature Vector-Based Similarity

If the input data consists of elements of a *normed vector space* one can use distance functions that exploit the space’s norm properties, such as the ones described below. In addition, one can use notions of relatedness that exploit the fact that it is possible to determine angles. By utilizing the correlation between distance function and similarity function, it is possible to derive similarity measures. However, if the input data does not consist of elements of a normed vector space, which is most often the case, e.g. when dealing with structured objects like graphs or strings, it is not possible to directly apply vector-space norms or angle determination. In those cases, it is necessary to transform the elements into elements of a normed vector space by applying a so-called *feature transformation*. This operation takes the attributes or features of the input elements that are most characteristic and thus most discriminating and creates vectors out of them. Choosing the right attributes is a non-trivial task in general and usually requires domain-specific knowledge. After the feature transformation, one can apply the distance measure common for vector spaces without regard of the actual data that the vectors represent. Figure 5-16 depicts the general idea of a feature transformation.

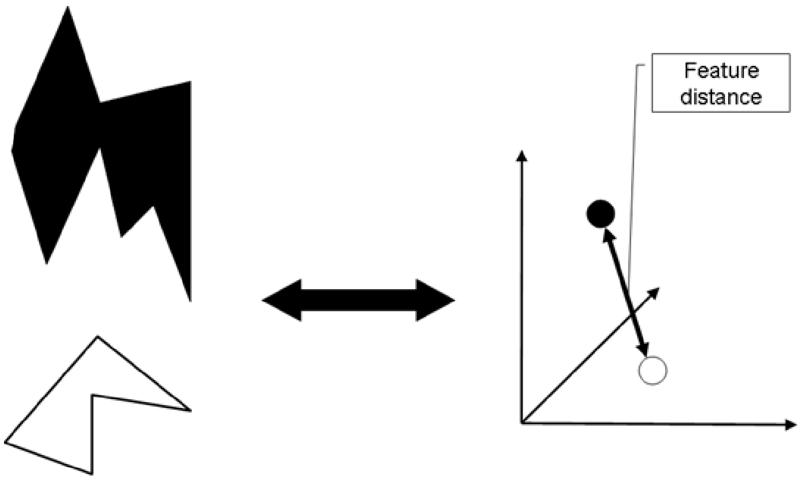


Figure 5-16 ***Illustration of feature transformation***
Source: Own Illustration

5.5.2.3.1 Vector Space Norm

A norm on a vector space can be seen as a measurement of vector lengths. It requires the norm function to be *definite*, *homogeneous* and to fulfill the *triangular inequality* as mathematically illustrated in the following equations:

$$\|x\| = 0 \Leftrightarrow x = 0 \ (\textit{definite})$$

$$\|\alpha \cdot x\| = |\alpha| \cdot \|x\| \ (\textit{homogenous})$$

$$\|x + y\| \leq \|x\| + \|y\| \ (\textit{triangular inequality})$$

A normed vector space is always also implicitly a metric vector space, since each norm can be used to generate a metric by using the following trivial equation.

$$\boldsymbol{d}(\boldsymbol{x_0}, \boldsymbol{x_1}) = \|\boldsymbol{x_0} - \boldsymbol{x_1}\|$$

The advantage of using vector norms as basis for metrics and thus similarity measures is their relatively low computation cost in comparison to metrics not using the properties of a vector space. The most common metrics applied in vector spaces are the so called *p-norms*, which are defined for each finite-dimensional normed vector space. In general a p-norm is defined as follows:

$$\|\boldsymbol{x}\|_{\boldsymbol{p}} := \left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} |\boldsymbol{x_i}|^{\boldsymbol{p}}\right)^{\boldsymbol{1}/\boldsymbol{p}}$$

and defined as follows in case of p converging to ∞:

$$\|\boldsymbol{x}\|_{\infty} := \max_{\boldsymbol{1 \leq i \leq n}} |\boldsymbol{x_i}|$$

where p denotes an integer equal to or greater than 1. In both cases $|x_i|$ denotes the absolute value of the ith component of vector x. Metrics that are derived using the p-norm are called Minkowski-metrics. The three most commonly used ones are the 1-norm, 2-norm and the maximum norm, which are shortly described in the following.

5.5.2.3.1.1 1-Norm

The *1-norm* is the sum of the absolute values of all components. In case of only two dimensional input data it is also known as the *taxi norm* or the *Manhattan norm*. This naming is due to the typical look of the norm in a two-dimensional coordinate system, that resembles a city block as one can find in e.g. Manhattan.

5.5.2.3.1.2 2-Norm

The *2-norm* is also known as the *Euclidean norm*. It represents the length measure that corresponds to the every-day expectation of what distance is. Thus in a two-dimensional space, using this metric normalized to 1 describes the unit circle, and in three dimensional space it would describe the unit sphere.

5.5.2.3.1.3 Maximum- Norm

The *maximum-norm* is also known as *Tschebyschew-norm*. It is a class of the general p-norm.

5.5.2.3.2 Cosine Coefficient

The *cosine coefficient* determines the cosine of the angle between two (feature) vectors. This measure is therefore only applicable to input data that is drawn from a vector space, or input data that has been transformed into a vector space using a feature transformation. If the angle is 0° the two vectors are considered to be equal and the similarity value is 1. An angle of 90° (π/2) will return a value of 0 thus indicating that the two vectors are not related at all. In case the angle is 180°(π), the resulting value will be -1 showing that the two vectors represent the exact inverse of each other. The calculation of the cosine coefficient is as follows:

$$\boldsymbol{cos}(\boldsymbol{\theta}) = \frac{\vec{\boldsymbol{A}} \cdot \vec{\boldsymbol{B}}}{\|\vec{\boldsymbol{A}}\| \, \|\vec{\boldsymbol{B}}\|}$$

Depending on the angle between the two vectors, negative values are possible. Values for similarity can however not be less than 0. Therefore, negative values are interpreted as an indication of no similarity. Only positive values of the cosine coefficient are considered. The formula for calculating similarity using the cosine function then becomes

$$\boldsymbol{sim_{cosine}} = \boldsymbol{max}(\boldsymbol{0}, \boldsymbol{cos}(\boldsymbol{\theta})) = \boldsymbol{max}\left(\boldsymbol{0}, \frac{\vec{\boldsymbol{A}} \cdot \vec{\boldsymbol{B}}}{\|\vec{\boldsymbol{A}}\| \, \|\vec{\boldsymbol{B}}\|}\right)$$

It is worth noting that the cosine coefficient does not take into account the length of the vectors, thus the frequency of a feature is not taken into account at all. It is often well-suited for sparse vector spaces.

5.5.2.4 Distance-Based Similarity

It is not always possible to find a suitable feature transformation for the input data. One type of input data for which this is true are unconstrained graphs, or even trees with unlimited size. Usually there is no way to find a feature transformation that is feasible and not too coarse. Often, one either chooses a feature transformation that contains too little information for a suitable discrimination of elements, or one is faced with a potentially infinite, but surely very high number of features, that are necessary to suitably discriminate elements. If no feature transformation can be found, it is therefore also not possible to use vector norms as the basis for a metric definition. In those cases, there is no other way but to define a metric or maybe only a distance function for the specific problem domain directly. Usually those metrics are more specific and more flexible for their respective problem domain. This is also the reason why sometimes *distance-based similarity* is used instead of a vector based similarity although it would be possible. The drawback is a typically much higher computational cost. The direct determination of similarity using a distance between objects is indicated in Figure 5-17.

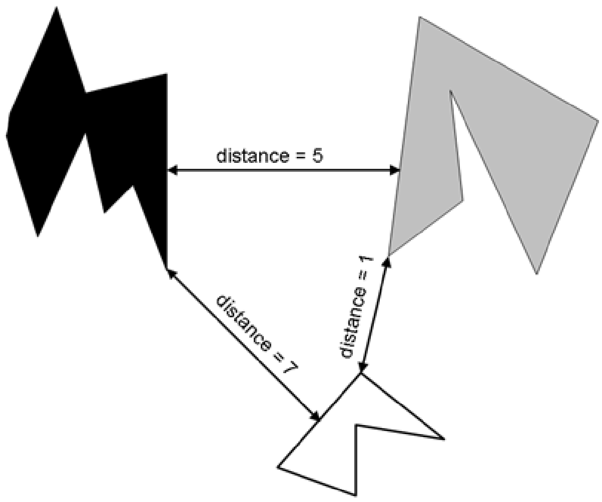


Figure 5-17 Illustration of the concept of distance-based similarity
Source: Own Illustration

5.5.2.5 Levels of Abstraction for Similarity Measures

Whenever similarity measures are used for supporting selections, an implicit assumption is made. One expects that the output parameter is related to the input entities that are used for similarity computation. For example, one assumes that the similarity of the task sequences that two individuals are currently engaged in, say procurement of high-value assets and procurement of high-volume, low cost materials, will be a good indicator for the perceived relevance of an interaction between those two individuals. However, if the result is relatively independent of the task sequences, the usage of similarity measures to support the envisioned goal might be less powerful independent of the algorithm used to determine the similarity. On the other hand, the level of abstraction can play an important role. In the example, it might be the case, that the similarity of task sequences is a bad indicator – it is a too broad concept – but the similarity of the single tasks that are part of sequence might very well be a good indicator. Two individuals that both have the task to assess the financial risk of a customer may rely on the same methods independent of the customer. In this case, the similarity in tasks is a good indicator. Hence, it is necessary to analyze which levels of abstraction can be relevant to understand which possible measures can be relevant. In the following, the possibilities always relating to this thesis' goal of relating two individuals and establishing communication in a way that is perceived beneficial for both, are discussed.

When defining similarity for individuals in a knowledge work context, one has to think about the desired *level of abstraction*. In general, similarity can be defined on the basis of *person attributes* that are reflected in or can be extracted from the system logs at hand. In addition, one can focus on the *previously executed tasks*. The tasks in turn can either be seen as "flat items" or as having further information items, i.e. *attributes*. Therefore, it is possible to model individuals in three increasingly concrete levels: firstly, only information about the individual itself is considered, secondly instead or additionally to the previous, the execution sequence of tasks is considered and thirdly additionally to the previous, the attributes of single tasks are considered. Figure 5-18 illustrates the three levels of abstraction graphically.

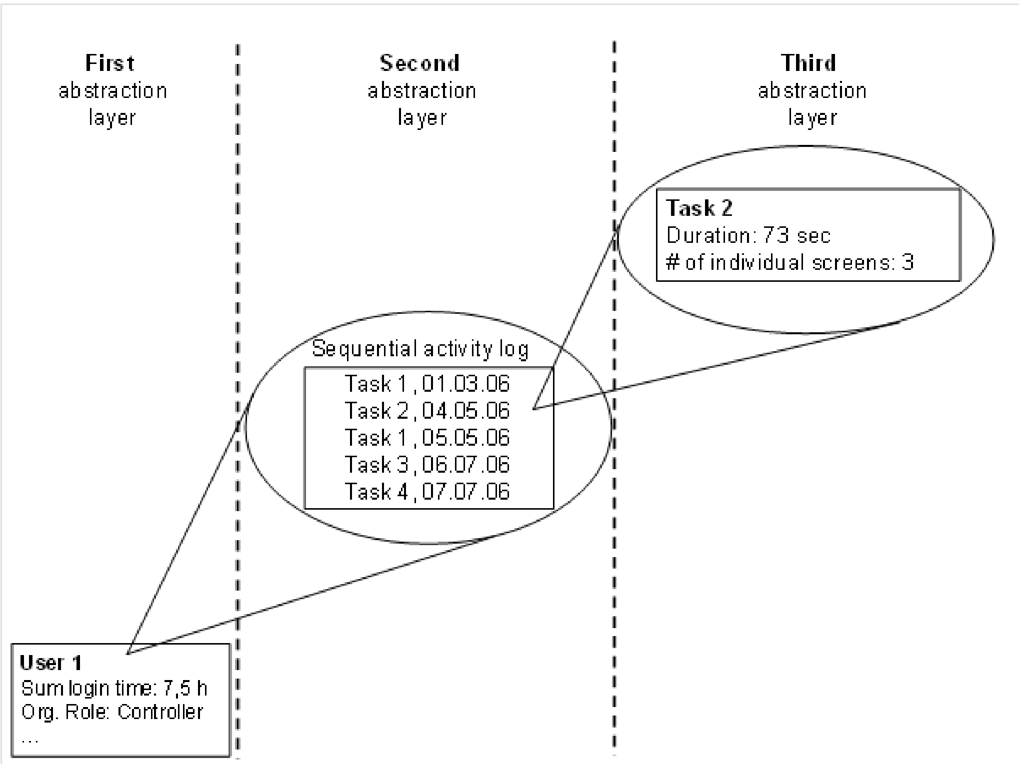


Figure 5-18 ***Different abstraction layers for user description***
Source: Own Illustration

An individual's attributes are used to store information about the individual itself, such as overall log in time in a certain application or the organizational role that this individual assumes. This meta-information about the individual can be used to compare individuals and thus finding e.g. coworkers that have the same organizational role.

Comparing individuals on the basis of their executed tasks builds upon the idea that a similar way of doing “things”, thus performing similar tasks, should reflect similarity in the kind of knowledge that an individual is interested in.

An individual's tasks usually contain additional information about the tasks. This can be used for a more concrete comparison. The idea is that tasks on their own do not convey enough information to be suitably compared, but e.g. the data that was worked on influences the actual meaning of a task. For example, one might want to consider a task *buy supply* with a contract volume of 100 € to be not very similar to the task *buy supply* with a contract volume of 1,000,000 €.

5.5.3 Keeping Effort for Users Low

One of the requirements for the Mediation Layer asked to keep the effort for all individuals involved in the overall process as low as possible. In this thesis' approach, there are three kinds of individuals: the knowledge seeker and individuals that are asked for help, with the latter consisting of those that do not want or cannot help and those that do (try to) help (see section 5.1.6). Capitalizing on the thoughts of Ye et al. (2008) and the discussions in section 5.1.6 about reducing the information overload as well in the section 5.5.1 about how to balance between contextual overlap and expertise, next the discussion focuses on describing how effort could be kept low. For this purpose, the following formula is adopted from Ye et al. (2008):

$$\boldsymbol{Effort} = \boldsymbol{1 * Asker_{find} + 1 * Asker_{ask} + q * N * Asker_{evaluate} + N * Recipient_{interupt} + p * N * Recipient_{skim} + q * N * Recipient_{answer}}$$

where $Asker_{find}$ is the effort of the knowledge seeker to find suitable interaction partners, $Asker_{ask}$ is the effort to formulate a proper request, $Asker_{evaluate}$ is the effort of the knowledge seeker to determine the quality of the response, $Recipient_{interupt}$ is the effort induced for all receivers of the message for the act of receiving a message, $Recipient_{skim}$ is the effort of a message receiver to skim the content, $Recipient_{answer}$ is the effort for a knowledge bearer to formulate an answer, N is the number of individuals that the request was sent to, p indicates which fraction of the receivers reads the message and q indicates the fraction of receivers that formulates an answer.

To reduce the effort there are various levers. First, one may reduce the cost of finding suitable interaction partners $Asker_{find}$. In this thesis' approach, suitable candidates are determined automatically using contextual overlap and levels of expertise as implicit input. Therefore, there is no effort at all for the user and $Asker_{find}$ is zero in this thesis' approach. There is one possible downside of a fully automatic approach. Yetim (2009) notes that users of recommendation systems may want to know the underlying reasons how the recommendation system came to the choice of the recommendation, in this thesis' case to the selection of possible interaction partners. In his contribution, Yetim talked about product recommendations. There, it is typically the case that users want a product but do not know which one to prefer. They do not know the exact features but they have some expectations or boundary conditions that should be met. When they are presented with a suggestion, they want to make sure that their expectations have been met and hence want to see why the selection was presented to them. On the other hand, if an individual faces a current challenge he searches help – a knowledge bearer that fits. In this case, it is more important to the knowledge seeker to get an indication of who to contact than knowing why it is this exact person that the system proposes. On top of this, often the knowledge demand cannot be formulated exactly and hence the evaluation of suitability is similarly complex for the knowledge seeker.

Another lever to decrease the overall effort is to reduce the effort for formulating a question $Asker_{ask}$. In this thesis' approach, this happens implicitly. Using the contextual overlap as one main criteria for determining interaction partners (see section 5.5.1), the knowledge seeker's request will be sent to other individuals that work on similar tasks and hence are more likely to understand the *same terminology*. This reduces the effort of the knowledge seeker to explicate his problem, as he may use the terms that are natural for him and may rely on some background knowledge of the receivers of the request. The same argumentation applies for the effort of determining the suitability of responses $Asker_{evaluate}$. Here again, since the potential recipients share contextual overlap with the knowledge seeker, they likely use familiar terminology to describe possible solutions. In addition, performing similar work increases the chance that they understand the problem not only generically, but in the context in which it appeared, which even more increases the quality and interpretability of the answers. The effort for skimming a request $Recipient_{skim}$ is likewise prone to be less, when the knowledge seeker uses familiar terminology.

While the effort associated with being interrupted cannot be affected, the number of affected individuals N can be reduced. Sending the request to many individuals increases the likelihood of receiving at least some responses. However, as Ye et al. note, " [..] the most significant waste of attention [in this thesis this is equal to effort; remark by the author] is that consumed by onlookers" (Ye et al. 2008, p. 5), which are those individuals that do not want or cannot respond to a request. The "lost effort" can be reduced, if the ones that do receive a request and are disturbed take a benefit out of the request. This is the case if they read the message and take value out of reading it, for example because they may have similar problems but have not asked for help yet or it may reinforce their understanding of the topic. This means the factor p of receivers who skim the message should be higher to increase overall benefit. The factor q, of receivers that answer should likewise be high because in this case the effort of skimming the message was not wasted but may lead to resolving the problem of the knowledge seeker. Additionally, the individual providing help may benefit himself. Rambow and Bromme (2000) reason that by trying to tackle a problem of someone else, in their case a layperson, the person supporting the problem realizes a resistance to what he knows that leads to a reframing of his expertise and a deeper understanding of the domain.

Yet, trying to limit the effort by connecting only those individuals that perform similar tasks, one has to live with trade-offs. On the one hand, there is the potential waste of effort for skimming over requests that an individual later does not answer. Schmidt (2000) notes that when it comes to collaborative efforts, participants are reluctant to care about the " [..] enormous contingencies and infinitely faceted practices of colleagues, unless these may impact on the our [sic] own work" (Schmidt 2000, p. 143). While this thesis' approach is not a collaborative effort according to the definition, as the interaction partners do not have a (obvious) shared goal, it is logical that the *information overload* that Schmidt addresses, holds for non-collaborative interactions as well. Possibly the argument applies even more since there is no shared goal that could justify the attention paid to others' doings. On the other hand, receiving messages that may not be relevant for an individual in the first place may turn out to spark new thoughts, an effect that has become known as *serendipity effect* (see for example Foster, Ford 2003). This effect is also used in open innovation approaches (Blohm et al. 2010) to find inspiring new ideas and in many social networking systems like xing.com and facebook.com to suggest connections between individuals. As a rule of thumb, serendipity effects are advantageous when creativity is of higher importance and less when problems need to be solved with a concrete relation to a certain topic. Therefore, in this thesis' approach, serendipity effects will not specifically be taken into account and there will be no support facility incorporated into the approach.

5.5.4 Summary

Summarizing the discussion about the *Mediation Layer*, one can conclude that a suitable representation of the users can be found in implicit models about the user. The implicit model consists of the tasks and their sequence plus possibly additional, application-specific payload data. It contains two aspects about the user, his expected *level of expertise* as evidenced by the number of times that the user has performed a certain task and the current and previous *task contexts* that the user is or was in. The discussion supported the proposition that only taking expertise as criteria for connecting knowledge seeker and knowledge bearer is in general not advantageous but contextual overlap should be taken into account as well. Building upon results from socio-psychology, the discussion brought up the argument that higher contextual overlap decreases the effort for formulating requests and answers and increases the likelihood of positive expectations towards the interaction. Furthermore, the four strategies that may be applied when both criteria – level of expertise and contextual overlap – are employed to find suitable collaboration partners were described. What's more, the discussion highlighted how using both aspects may reduce the effort for all affected individuals. By only sending requests to individuals that are likely and capable of answering, the "waste" of effort for others is limited. The increased overlap in context decreases the effort for the interaction. Explicitly taking the user's current context as part of the "automatically compiled query" for suitable interaction partners further ensures that interactions are more beneficial. Results from socio-psychology underline this, as work expertise is less often generalized but crystallizes around the work context. This sub-chapter also discussed the definition of *similarity* as used in this thesis. Different kinds of similarity measures were introduced, describing the difference between *feature vector-based similarity measures* and *distance-based similarity measures*. Moreover, the argumentation brought up that similarity measures may operate on different *levels of abstraction* with respect to the tasks performed by individuals, namely information about the individual, information about tasks executed by the individual and finally also attributes of those tasks at the highest level of detail.

5.6 Data Acquisition Layer

This thesis' approach aims at supporting the execution of knowledge-intensive processes that are in turn supported by IS. As addressed before in this chapter, a viable way to do so, is by tapping into the traces that knowledge workers leave when they work with IS. Consequently, coming from the black-box view on the *Data Acquisition Layer* in section 5.3.2, the following functionalities that the layer should support have been carved out:

- Putting minimal or no burden on the user
- Usage of suitable data entities for expertise determination
- Coping with frequent changes in the model of users
- Employment of knowledge structures that the users are acquainted to
- Usage of data from operative IS
- Ability to integrate more than one data source

The requirement to put minimal burden on the user, calls for an automatic feature extraction that ideally needs no feedback from the user. To enable this, as a first step, the kinds of IS that may be used in concert with this thesis' approach are analyzed. Then the focus shifts to the data entities that may be used to suitably determine expertise. The additional requirements such as the employment of knowledge structures are discussed subsequently, along with the process of extracting and representing the data.

5.6.1 Data Sources

As a first step, it is necessary to investigate which IS can be tapped into. Seeing that the features of ERS are to be mimicked, it makes sense to analyze their use of data sources for the elicitation of expertise. Section 3.3.2.1 established the assumption that ERS typically rely on one of the following data sources:

- Communication-related sources (e.g. e-mails, newsgroup postings)
- Document-related sources (e.g. files, documents in a CMS, websites)
- Appropriated Knowledge Assets (e.g. citations)
- Interaction recordings (e.g. resource usage, program usage)

There is a conceptual difference between the first three sources (communication-related, document-related and appropriated assets) and the last one (interaction recordings). The first three sources contain, often textual, creations of their authors (or senders respectively) *about* something. The interaction recordings on the other hand contain information about the *execution* of something. This relates to different kinds of knowledge as Lundvall and Johnson (1994) note. In the first three sources, one will likely find *know-what* – knowledge about facts – and possibly *know-why* – knowledge about rules or (logical) connections that explain the inner workings of a phenomenon. From interaction recordings, one may on the other hand find evidence for *know-how* – knowledge about how to do things or "practical" knowledge as the interactions are evidence for actually performing a task. Independent of the source, ERS try to foster the fourth of Lundvall and Johnson's (1994) identified type of knowledge: *know-who* – knowledge about who knows what. Jensen et al. (2007) found that the differentiation between know-what / know-why and know-how also induces different mechanisms for learning and innovation. They note that "[..] one mode is based on the production and use of codified scientific and technical knowledge [know-what / know-why; note form the author], the Science, Technology and Innovation (STI) mode, and one is an experienced-based mode of learning [know-how; note form the author] based on Doing, Using and Interacting (DUI-mode)." (Jensen et al. 2007, p. 1). This thesis' approach wants to support the users in interacting and exchanging knowledge and experience, hence interaction recordings seem most promising. This is supported by Seid and Kobsa as they indicate that if finding an expert to use him as gateway to or instead of an information record of some kind, it is justified to use the universe of this information record to extract expertise assumptions (Seid, Kobsa 2003, p. 13). The focus of the contribution at hand lies on connecting individuals that share context and that exchange knowledge about tasks that they perform and hence following the argumentation of Seid and Kobsa, interactions recordings are a suitable means for this purpose.

Referring to the previous discussion about IS that are used by knowledge workers allows to draw some conclusion. While knowledge workers may use arbitrary IS in an ad-hoc fashion during their work, the results of their work are either documents that are exchanged through some medium, like e-mail or the results are stored in some IS that supports the enactment of tasks. Focusing on a system's interaction logs, one may interpret or at least relate entries in the logs with the tasks that have been executed and then the logs contain sequences of tasks. In this case, there appears to be a relation to processes that are also sequences of interactions. However, there does not necessarily have to be an explicit process. Dumas et al. (2005) researched systems that in some way or another support processes be it explicit or implicit. Hence, the same class of systems that Dumas et al. investigated may be a viable source for this thesis' approach. Dumas et al. (2005) call IS that rely on process models *Process-Aware Information Systems (PAIS)* independent of whether the models are explicit or implicit. They distinguish four different subgroups of those systems that differ in the strictness of process model enforcement.

The first group - *Workflow Systems* - has a strictly enforced process model and systems of this kind are designed for high volume process enactment. There may not be any deviation from the model at all. Only alternatives that are in the model can be executed. The second group - *Case Handling Systems* - usually does support a loose form of a process model, but the user can choose to deviate from it slightly. Those systems were originally designed to support knowledge workers, as they cater for flexible, self-chosen enactment sequences. The third group - *Ad-Hoc Workflow Systems* - allows the user to completely deviate from a potentially present process model. The models are just reference models and are not normative. The final group - *Groupware Systems* - does not have a process model, but rather offers a structured way of e.g. storing and retrieving information. Any of these systems may be a viable source for this thesis' approach.

5.6.2 Data Entities

Since the kinds of IS have been addressed sufficiently, in a next step, it is necessary to define what kind of data entities are to be extracted from these IS. The intention is to use the tasks that an individual has executed, i.e. the sequence of tasks, to determine the *contextual overlap* with other individuals. However, there may be different data stored during the execution of a sequence of tasks. In addition, there are various ways to define context. One way to classify different elements that can be used to describe a context can be found in (Shkundina, Schwarz 2005).While analyzing how to support automatic provision of documents in workflow systems, Shkundina and Schwarz (2005) determined the following six contextual elements:

- *Informational aspect*: Documents that have been used or created during the execution of a task

- *Organizational aspect*: Persons, roles or organizational units that are involved with the task
- *Behavioral aspects*: Atomic interactions on low levels like “print document” or “add bookmark” and their sequence
- *Operational aspect*: Other applications that are being used during the execution of the task
- *Causal aspect*: Goals of the task as well as of the user
- *Chronological aspect*: The order of additions and removals of contextual elements during task execution

The *informational aspect* seems to integrate nicely with the way that many ERS work – they use text-based documents for information retrieval about possible expertise. However, to use this contextual element, the documents have to have an application-specific connection to the task, e.g. in a CRM system, the creation of a textual description of a customer is by definition assigned with the task “describe customer”. Another possibility would be to observe the user’s utilized documents while he is executing a task in the application and then assigning the created document with the current task. This may however be a challenging undertaking, because it assumes that, even when a knowledge worker is engaged in parallel tasks, it is possible to deduce which documents belong to which task.

In terms of the *organizational aspect*, the executor of a task certainly is relevant context information, especially in the setting of this thesis where connecting those individuals and hence executors of tasks that have a sufficient contextual overlap is the predominant goal. However, larger entities such as organizational units play a smaller role in this thesis’ approach that aims at actively crossing intra-organizational boundaries to enable knowledge transfer according to task not to affiliation within the company.

The *behavioral aspect* that concerns itself with the atomic interactions of users with the supporting IS, has also been investigated in other research projects (Rath et al. 2006; Kröll et al.). There, low-level events like mouse clicks and key strokes are logged using an installation into the operating system of the user. The events are then aggregated into event blocks. For example, a mouse movement, a click on the left mouse button and a subsequent opening of a window can be an event block “open program”. Event blocks are then aggregated into tasks that in turn belong to a process instance. Analogously, in ERS, the use of low-level events or information like the time spend reading documents or operations in spreadsheets has been proposed as input for adaptation of recommender systems (Terveen, Hill 2002). Granitzer et al. (2009), however report that, while the approach described above works with moderate accuracy in their use case, it was only possible to correctly deduce a task from the interactions of an individual in an average of 77 %. Further, they report that generation of processes from tasks using process mining techniques failed in their use case due to too little accuracy of the result. Owing to these challenges and the inevitable need for considerable degrees of user feedback and hence effort on the user side for training the classification mechanism, this way of support is not adopted as contextual element into the approach described in this thesis.

The *operational aspect* seems appropriate for this thesis' approach. Overlap in context may not only be defined exclusively by the task that a knowledge seeker needs help for. Looking at the parallel and possibly interwoven execution of similar tasks instead of only one task may yield a more accurate determination of contextual overlap between two individuals. For example, if some individual asks for help while executing a task "create new sales order" while he is also performing the task "create new customer" the two hypothetical individuals that are also currently performing (or have recently performed) the task "create new sales order" appear equally suitable for contacting. However, if one of the two is also currently performing the task "create new customer" like the individual in need of help and the other one is instead additionally performing the an additional task "evaluate sales history", it becomes obvious that the first candidate has higher contextual overlap than the second one.

Causal aspects of context, i.e. the information about what an individual's goals are would be most beneficial as then this thesis' approach could connect those individuals that have the same goal or a similar goal. However, while there are some ideas how to support this task (Schwarz, Roth-Berghofer 2003) the elicitation of goals is non-trivial and has to fight against high levels of uncertainty (Shkundina, Schwarz 2005). Due to these challenges and the inevitable need for considerable degrees of user feedback and hence effort on the user side for training the classification mechanism, this thesis' approach chooses to neglect this of contextual element.

The order in which tasks have been performed is logically connected to know-how, i.e. knowing in which order to best do tasks to reach one's goals. Therefore, the *chronological aspect* of context makes sense in this thesis' approach.

In essence, *documents* that are related to a task may be suitable for the description of contextual overlap as well as expertise. However, this applies only if there is a reliable relation between the two that does not have to be reconstructed with means that increase uncertainty. The *executor* of a task is naturally relevant for this thesis' approach. Further, operational aspects may be relevant when they indicate *other tasks* that are being executed. Finally, information on the *chronological aspects*, i.e. what has been done in which order, is relevant for this thesis' approach. Rephrased in colloquial words: The essence for determining contextual overlap and expertise will be "what has an individual done, how and in which order". The "what" is the task that has been performed and also the tasks that have been performed and additionally describe the current contextual setting. Depending on the concrete IS that the data entities come from, this could be an application function, a transaction or a program or screen of a program that is called. The "how" is very much dependant on the IS. While some systems may store quantitative information about tasks that can be used to deduce performance, this is not true for all systems. In the most generic interpretation the "how" part may be interpreted broadly to just indicate that a task has been performed (somehow). The "in which order" part implies that there is a temporal order in the log files that can be used to deduce the interwoven or sequential order of tasks, i.e. the structure of execution.

Looking at other areas of research, the problem statement is similar to the field of *process mining* (c.f. Agrawal et al. 1998), where entries in log files of various PAIS are used to reconstruct a process model. The de-facto standard, open-source support tool for this task, the ProM framework (van Dongen et al. 2005), uses a basic data entity, the *AuditTrailEntry*, that is designed to be compatible with every log of tasks. It encompasses the *originator* (= the individual executing a task), a *timestamp* (= the time when a task was started), an *event type*, that is optional and indicates whether a task was started, suspended, resumed, etc., a *WorkflowModelElement* (= the task that was executed) and optional, arbitrary *payload data*. This is essentially the same information that the discussion before yielded to adequately describe context for this thesis' purpose.

Using the context information described above as evidence for expertise, one can assume higher expertise with a task for those individuals that have executed this task many times. In analogy to Terveen et al. (2002) who discussed the usage of number of times a topic was discussed as evidence for expertise about this topic, it should be noted that performing a task often does not necessarily show expertise but it is a valid assumption. In addition, the availability of sequences of tasks may further support means to deduce expertise. For example, there are some ERS that used web browsing sequences to deduce expertise on topics of these website. Seid and Kobsa noted that " [...] one may question their basic assumptions on the ground that browsing shows more of the interest of someone rather than expertise." (Seid, Kobsa 2003, p. 8). In this thesis' approach, however, the focus is on sequences of operational tasks and those are typically not explored let alone explored for the mere interest but actively used for operative tasks, which remedies this challenge of the assumption in this thesis' setting.

The data entities as described above – information about task, time of execution and executor along with possibly available payload data –therefore appear suitable as they convey context in a meaningful way and may act as evidence for expertise at the same time.

5.6.3 Extraction and Representation of Data

The discussion so far revealed what kind of data is eligible for the use in concert with this thesis' KM approach. As a next step, it is necessary to investigate how to extract this data and how to represent it for further processing. The relevant data is stored in the *logs of operative*

IS. IS typically log users' activities for administrative reasons e.g. authorization control or for analytical purposes e.g. web server logs are used to determine (aggregated) usage behavior.

In terms of architecture, for the extraction one has to keep in mind some of the requirements deduced before. First, there may be frequent changes to the implicit user model. As a matter of fact, using *task sequences* as the fundamental data entity, there will be an update of the implicit model of context and expertise that is resembled in the structure of the executed tasks, whenever an individual starts a new task, and hence many times a day. This can create a technical problem as the extraction of data puts stress on the operative systems. Additionally, the implicit model may span an extended period of time, depending on how far back the interaction sequences should be considered. For expertise elicitation a larger time window may be suitable, i.e. it is relevant to know how often an individual executed a certain task in say the last year. Yet, for context determination, the time span may be much shorter, i.e. it is less relevant to know whether two individuals did similar tasks in the last year than knowing if they did similar tasks in the last minutes or hours as this better reflects contextual overlap. Therefore, in order to be flexible, a suitable architecture implementing this thesis' approach should store the implicit model and hence the relevant parts of the application log files in a separate system. In this separate system, the level of expertise may be calculated using data that covers an extended period of time. In addition, the determination of context may be pre-calculated there. The more recent part of the context model can then be adapted with real-time or near real-time information about the user's interactions from the operative IS to reach an up-to-date model of the work context. This way, frequent updates are possible without putting too much load on operative systems or their log files and context determination stays up-to-date.

Further, the investigation of requirements indicated that relevant data may be distributed across various IS. For example, a knowledge worker may enter a complex sales order for a customer in an ERP system, but before that, he checks the sales history of this customer in an analytical application of a customer relationship management system to determine which price structure to apply for this customer. The two tasks "check sales history" and "enter sales order" are part of this individual's (implicit) process, but are spread across two different systems. To be able to analyze the process anyway, different technical obstacles need to be addressed. First, there needs to be one system where information from the logs of both IS can be send to, in order to later combine the information. Second, the data formats need to be harmonized, to the format described above, hence there needs to be information about who did what when, and possibly how (= application-specific payload data). Third, the technical users of the two systems need to be related to relate to the one real-life user. If a central authorization system like an LDAP exists, this is easier, otherwise there need to be means to harmonize possibly different user names of the same individual.

Finally, the requirements analysis showed that it is necessary to include item names that the users are acquainted to when building the implicit context and expertise model. This is relevant later in the *Representation Layer*, where structuring of interactions is relevant and structuring according to tasks or task taxonomies using the “normal” language of the user is advantageous. For example, a log file may have an entry “cso,user1,20110101”, which is hard to interpret for end users. The user would rather expect the information “The user number 1 has performed task ‘create sales order’ on 01.01.2011”, i.e. the language the end user is used to instead of technical terms.

5.6.4 Summary

Summarizing the discussion about the *Data Acquisition Layer*, one can concluded that any operative IS that is a process-aware information system (*PAIS*), i.e. that has an explicit or implicit process model, is compatible with this thesis’ KM approach. The data that needs to be extracted from those systems or more specifically from the logs of these systems needs to contain what has been done (*task*), who has done the task (the *knowledge workers*), when the task has been done (for recreation of *sequences*) and possibly *payload data* that is specific for the application. This is sufficient information to deduce *expertise* with the execution of the tasks and to determine *contextual overlap*, which is addressed in the *Mediation Layer*. Further, in terms of architecture of an extraction mechanism the discussion showed that there should be an (external) single system where the information contained in the logs of one or more IS can be sent to for long term storage. Information about recently executed tasks, on the other hand, needs to be forwarded to the *Mediation Layer* immediately. Finally, the data extraction mechanism should be able to create representation of the log information that is natural for the end users, as it will be used later in the *Representation Layer*. The approach as described fulfills all the requirements that were determined in section 5.3.3. The extraction from log files works without putting burden on the user. The extracted data entities were shown to reflect context and expertise suitably well. Further, the section explained how to deal with frequent changes of the model and how to integrate different IS into the approach. The data is extracted from operative IS and hence reflects the task context well. Addressing the need to use terminology that users are acquainted to fulfills the last requirement that was determined.

Much of what has been discussed in this section is specific to the application that hosts the implementation of this thesis’ approach. Nevertheless, chapter 6 later analyzes which IT artifacts can be created in a sufficiently generic way to support the extraction of data from arbitrary IS. Section 6.3 will also discuss a data model that suits the requirements as specified in this section and that is able to transform data from the logs of IS into terms that are understandable by end users.

5.7 Summary

This chapter described the design of this thesis' central artifact – the approach for fostering knowledge exchange based on relevant activities. This way it supports the "[..] move from the metaphor of *knowledge management* to a new metaphor, *expertise sharing*, which promotes focusing on the inherently collaborative and social nature of the problem." (Ackerman, Halverson 2004, p. 1).

The artifact design started by collecting preliminary design considerations that are informed by pragmatic needs. Building upon the analysis of contemporary approaches supporting the inter-personal knowledge transfer in chapter 3, the idea was born to utilize a *combination of ERS and knowledge networks* to arrive at a viable solution. To combine the two approaches, in this chapter it was shown that *task context overlap* is an appropriate "glue". Subsequently, the kinds of tasks that a knowledge worker might perform were categorized. It turned out that those that are performed in or supported by *Process-Aware Information Systems (PAIS)* (Dumas et al. 2005) are suitable for this thesis' approach as they are observable, interpretable and typically represent the final or a fundamental step of a knowledge worker's efforts. Then, the discussion focused on the influence of *anonymity* on individuals using this thesis' approach. The argumentation concluded that both, relative anonymity and the lack of it may be beneficial depending on the use case. Acknowledging the fact that functionality that has been researched in the area of ERS is adopted in this thesis' approach, the problems inherent to them were analyzed. In addition, a description of how this thesis' approach may *address these problems in expert determination* followed. Furthermore, two categorizations of *design dimensions for expert recommender systems* were utilized to integrate this thesis' contribution into the current state of the research and to better understand which additional aspects have to be take into account. In a subsequent step, the necessity to *embed the approach* into operational IS was elicited and discussed. The investigation of preliminary design considerations concluded with a thorough look at risks of information overload, the associated costs and how this thesis' approach can limit them.

As a next step, acknowledging the principles of *theory-driven design* (Briggs 2006), two theoretical models that inform the design of this thesis' approach and that represent a solid basis were examined. The first is *Social Capital Theory (SCT) applied in the context of intellectual capital building* (Nahapiet, Ghoshal 1998), that sets out to describe what effects the levels of knowledge contribution. The results of two instantiations of the model and their empirical results (Wasko, Faraj 2005 and Law, Chang 2008) were investigated to inform this thesis' design. The second theoretical model examined in this chapter also inspects what effects the levels of knowledge contribution (Olivera et al. 2008), but uses another foundation of theories from psychology: theories of *problem solving* (Newell et al. 1972) and *cognitive motivation* theories (Kanfer 1990).

In a subsequent step a *black-box view* onto the building blocks of the approach was given. Accumulating the *influencing factors* for the design of the approach posed sufficient information about which requirements need to be fulfilled. More specifically, the identified drawbacks and the conclusions drawn in chapter 3, the factors that generally determine success of KM approaches and KMS-supported approaches in particular in chapter 4 and the preliminary design considerations of section 5.1 were aggregated as influencing factors. This resulted in a design with three layers of concern – a *Representation Layer* that deals with the “visible” part of the approach, a *Mediation Layer* that resembles the functionality of ERS and that fulfills the task of matching users and a *Data Acquisition Layer* that deals with acquiring the necessary data for the support of the approach from operative IS. The discussion of requirements and how to address them also showed that some influencing factors are implementation-specific and need to be aggregated into the collective construct *organizational context* indicating that real-life implementations will have to treat those according to their context.

To further detail the approach, each layer was then described in elaborate detail starting with the *Representation Layer*. Here, one can distinguish functionalities that support *network size management*, general *management of the network* that is created and functionalities to support *user interaction*. The first class of functionality – network size management – builds upon task context similarity using it as the crystallization mechanism for new networks, as suitable (meta-) information to find existing networks and as helpful mechanism to separate too large networks into meaningful smaller sub-networks. The second class of functionality – general management of the network – dealt with the perspective of an operator. Here, questions such as how learning effects between two individuals may best be spread to a larger group are relevant and it is detailed how this thesis’ approach can support this requirement. Also aspects like activity measurement of the resulting network, different roles that individuals in the network may have and methods for incentivation of activity were discussed. Additionally the discussion focused on the effects of anonymity and how to make situation-specific use thereof. The third class of functionality – support of user interaction –describes functionalities that affect the users’ experience while benefiting from the approach. Here, the discussion focuses on how a user’s context is taken into account during the representation of interactions and how integration with operational IS should be addressed. Then, structuring of interactions within the network is talked about and the question how users may guide their personal content views is addressed. Finally, possible means of the Representation Layer to increase the likelihood of contributions are determined and elaborated on.

The design of the second layer of this thesis’ approach – the *Mediation Layer* – then guided the discussion. It supports the context-dependant creation of connections between individuals building upon suitable user models. As a first part of this layer, the question what a suitable representation would be, needed to be answered taking into account the two dimensions *con-*

textual overlap and *level of expertise*. Results from psychology about expert-layperson interaction, inter-personal attraction, group formation and trust (Ziegler, Golbeck 2007; Homans 1958; Byrne 1971; Zimbardo 1983; Bromme et al. 2004; Clark 2007) informed the setup of the implicit user model. Subsequently, the term “similar” was closely inspected and how it can be interpreted in this thesis’ approach. To this end, different *similarity measures* and distance measures were inspected and how they operate on task contexts. Finally, the Mediation Layer’s ability to contribute to *reduced effort for individuals* was pointed out inspecting each “item” that creates effort along the interaction process between individuals and describing how the approach can reduce or remove them.

The last section of this chapter described the *Data Acquisition Layer*, which deals with acquiring the necessary elements to create the user model in the *Mediation Layer*. The investigation first focused on the kinds of IS that may be valuable *data sources* for the extraction of necessary data concluding that any PAIS, i.e. system with implicit or explicit process model, can be used as data source. Then a description of the elements of data that could be used for the determination of context followed, taking into account different definitions of context. The argumentation concluded that as minimum criterion eligible *data entities* contain a task (name), a task executor, a timestamp and optionally additional payload data. Finally, the question how the *extraction* and subsequent (internal) *representation of the data* should be tailored is addressed, taking into account that there should be no user effort, the data should be up-to-date and stored with information that is understandable in the *Representation Layer*.

With the exception of the influencing factors that are determined by the organizational context, the described model fulfills all requirements that were deduced in section 5.3. While the concrete implementation of the approach will be specific to its application context, there is some functionality that is generic enough so that it may be integrated into a *supportive framework* to help adopters of this thesis’ approach. This will be the topic of the next chapter which describes this framework that supports the core aspect of this thesis’ approach: the use of *similarity* in the determination of contextual overlap.

6 Determining a Suitable Similarity Measure – A Supportive Framework

"Ein Mann, der recht zu wirken denkt, muss auf das beste Werkzeug halten."
Freely translated: "A man who wants to do things right, must rely on the best tools."
Johann Wolfgang Goethe (1749-1832), German Poet

The previous chapter elaborated on this thesis' expert mediation approach. It relies on connecting knowledge workers that benefit most from mutual interaction. A beneficial fit is determined by connecting those knowledge workers that share sufficient *common context* to be able to talk to each other in a shared language. In addition, the selection ensures that they have an *overlap in the tasks* that they worked on so that the interaction partners are most likely to extend their knowledge in the field of applications important to them. However, two interaction partners will never have the exact same context, i.e. the exact same work behavior in terms of order of tasks, personal background, and hence processing speed, etc. Therefore, the matching of suitable interaction partners needs to adhere to a *best fit*, which are those individuals that have the most in common concerning their interaction. It boils down to finding interaction partners that are *most similar*.

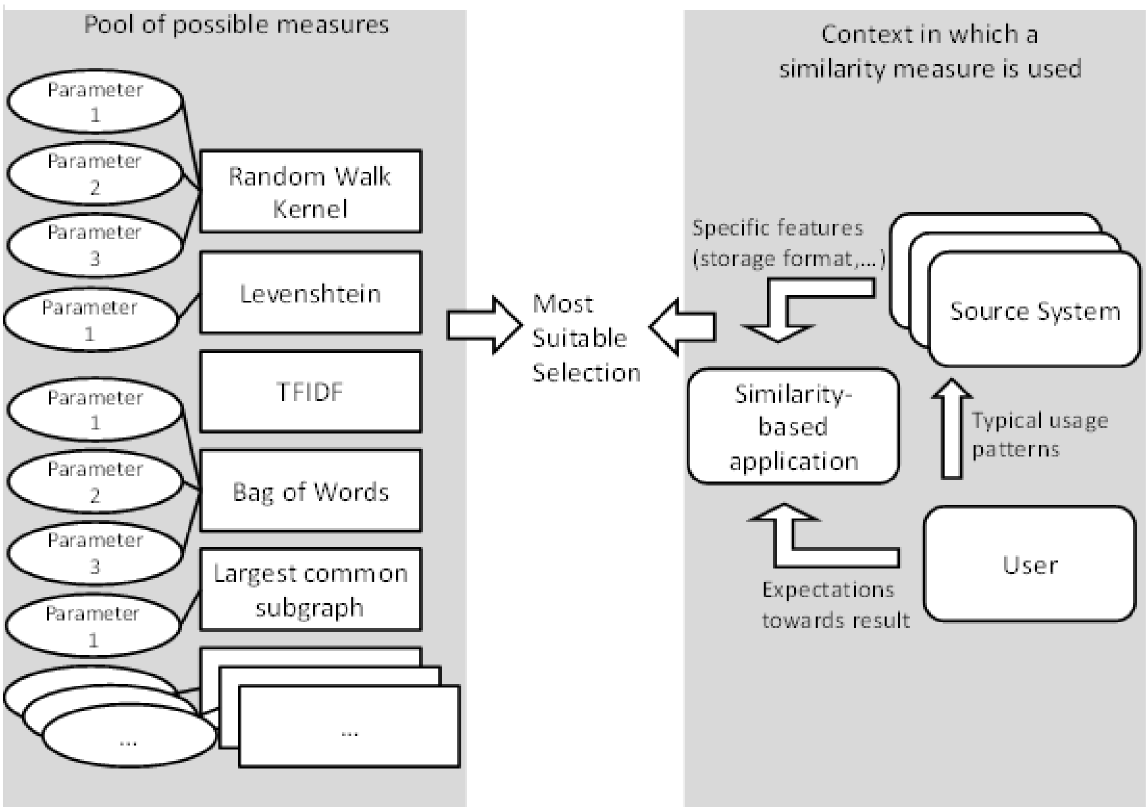


Figure 6-1 ***Requirements for suitable similarity measure in relation to pool of possible measures***
Source: Own Illustration

Similarity is, however, not a concept that conveys meaning by itself. It rather needs to be seen in the context of an application that relies on it and other factors as also illustrated in Figure 6-1. The context depends not only on the kind of IS that utilizes an appropriate notion of similarity, i.e. the goal that is pursuit – in the context of this thesis: expert mediation – but also on further contextual determinants. The *source systems* that are used for the later use case are important as well because they might pose specific requirements inherent to their storage format or interaction patterns. In addition, the way in which users of the source system typically work and their *expectation of what is a suitable result* for a computed expert mediation may vary significantly. On the other hand, there is a large pool of *possible similarity measures* (see also section 5.5.2.2) to choose from and to increase the scope, each of them typically can be configured to sometimes extensive degree. This is why finding the most suitable similarity measure for a certain context is far from trivial.

While different applications that rely on similarity will have different requirements, the determination of similarity must not necessarily change with respect to the algorithms, but only with respect to their parameterization. This supports the idea to create a general solution, i.e. *a framework*. The requirements of this framework will be addressed in the first section of this chapter. A framework for supporting such applications will have to cope with varying input *data sources*, must anticipate that it might be necessary to *label the input data* and must support the potentially necessary *generation of process models*. Furthermore, it must offer a considerable amount of *different similarity measures* and must offer an interface for applications that want to use the framework. Finally, support for *automatic tuning of parameters* is a useful feature to support.

A suitable supportive framework needs a suitable data format. Promising candidates need to be investigated and the most suitable chosen. In this chapter two promising candidates will be inspected and *MXML* will be selected as internal data format of choice. Having identified the natural occurrence of two different storage paradigms, *event-driven* and *state-driven*, and the need to support the event-driven format in this thesis' solution, a *transformation language* needs to be thought of. Consequently, a powerful transformation language will be developed in the course of this chapter.

Structural properties of executions may be valid sources for similarity determination. However, what is reflected in logs is only a linear sequence. Building upon a model, that is possibly reconstructed using *process mining* algorithms, this chapter shows how the structural properties can be uncovered. To support the process mining efforts the integration of an existing framework, ProM, into the framework that is developed in this chapter is described. In addition, the *transformation* of event-process chains and petri-nets into simple graphs that are the necessary input format for many similarity algorithms is elaborated on. The discussion includes argumentation for why this transformation is adequate.

After having conceptualized the framework it was implemented in Java fulfilling all elicited requirements. The import is supported by utilizing the ProMImport framework, while the mapping of data is supported by a *self-defined Extensible Markup Language (XML) application*. Classification is supported through an interface for list formatted labeling information. The process mining module builds on ProM, a renowned framework in the field of process mining. 14 configurable similarity measures are offered by the framework, that future applications can use through a clean interface. Additionally the applications can utilize the framework for *automatically generating* results of different *parameterizations*. All modules are designed in a way that allows easy extension with additional implementation variations.

6.1 Requirements for Appropriate Tool Support

Here, the *requirements* for a framework to support the selection and configuration of measures for similarity-based applications are documented. Requirements engineering can be done in many different ways (Kotonya, Sommerville 1998). Sources of requirements can be for example domain knowledge, existing systems, users, standards or regulations. As the main goal is to facilitate the determination and selection of similarity measures for a broad range of systems in the following existing frameworks, source systems, data format standards and applications that use similarity measures are used as source for requirements elicitation. Each requirement is annotated with a number for reference in the outline of the framework's architecture later.

A similarity measure is appropriate if it supports the goals of the target application. Therefore, the determination of an appropriate similarity measure consists of two phases: first, *selecting and configuring* a similarity measure and second, *checking the fitness* of the measure for the target application. However, a prerequisite is *having data* that is suitable to act as source for similarity information. This step is particularly important, because the selection itself greatly influences the result of the similarity determination (Guy et al. 2010). Therefore, a generic framework must have the three components as shown in Figure 6-2. In the following, each identified requirement references its corresponding component as indicated in this figure.

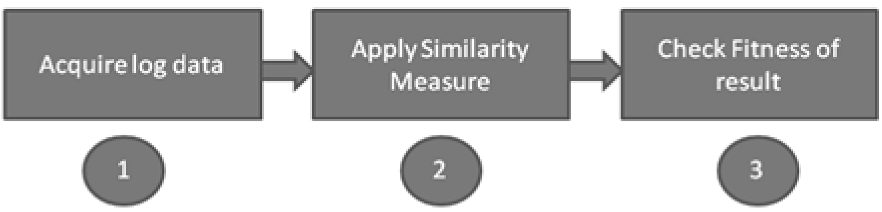


Figure 6-2 ***Steps for determining suitable similarity measure***
Source: (Schmidl et al. 2011b)

The requirements for the first component are elicited by inspecting different process PAIS (Dumas et al. 2005), such as ERP systems, project management systems and personnel management systems, the format of their data and how the log information is interpreted by the application and by persons.

Whenever it is desirable for an application to utilize the similarity of task sequences, the first step is to retrieve those task sequences from source application logs. However, many applications may serve as sources for information and their *data formats* also may take many forms. Some applications, for example ERP systems, store log information in databases while others such as web servers use files for this purpose, which are accessed differently. A requirement for the framework, therefore, is:

> The framework should support both – data stored in files and in databases (*R.1a*)

In addition, the *format of the data* may vary. While data in a database is structured by definition, file-based logs can be stored in comma separated value (CSV) files, using plain text with or without providing header information about the meaning of each section in a line, which is a still quite common solution. Other file-based logs are stored using some XML dialect with or without providing a schema definition along with it, especially when interoperability is important. Yet other IS use a log structure that is not plain text, is proprietary and needs parts of the IS logic to decode the log. Since all of these data formats are found in applications that can act as data source, the framework should fulfill the following requirement:

> The framework should support the retrieval of data from sources that are structured, semi-structured or use structures in a proprietary format (*R.1b*)

Additionally, the *granularity of the log information* may differ. Some applications log every user interaction, such as web servers, while others only log certain events including, for example, the change of a status indicator in a project management system. The same is true for *context information* that goes along with the log entry, which also can differ significantly in its extent. For example, while browsing in an intranet, much context information of a user is typically at hand, while anonymous access to an internet site offers less context information. Therefore, another requirement for the framework is as follows:

> The framework should be flexible enough to handle both rich data sources and to extract or amend less rich data sources (*R.1c*)

IS also differ in their *pervasiveness*. Some IS log user interaction in the background with little or no user involvement, such as web servers while users browse through a website, while other IS only write into their logs when explicitly requested by the user, such as for example accounting systems. This influences the granularity and the possibilities for interpreting the log, because in the first case one often needs to interpret implicit behavior while in the second case the intention of a user is more explicit and related more strongly to the log entry. Depending on the knowledge about the process that is supported by the IS, it is possible to amend log data with context information. This creates another requirement for the framework:

> The framework needs to be agnostic to how the data is captured from a technical point of view, but needs to provide means for amending the data with implicit information (*R.1d*)

Also, the *kinds of stored data* differ. Some applications store an event, or task respectively, in their logs, i.e. what has happened. Others store data that reflects the situation after a task has been performed, i.e. the result of what has happened. A web server for example might store the event “page index.html has been requested”, while a project management system might store the status “project budget is (now) 100k €”, but not the event itself that increased the budget to this amount. Another requirement for the framework, therefore, is as follows:

> The framework should have the capability to transform log information containing status snapshots into log information containing status changes (*R.1e*)

Independent from the characteristics of the data source itself, more than one application log may contain information relevant for a task sequence, i.e. the information contained in one application log can *augment information* from another. For example, a project management system could contain the execution history of a project, while in a separate accounting system, information about consumed budget is kept. This is why the framework should fulfill another requirement:

> The framework should allow the flexible and iterative enrichment of log data from multiple sources (*R.1f*)

Being able to import data from arbitrary data sources and being able to transform it into a suitable format, forms the basis for the second component of the framework in Figure 6-2. It allows applying similarity measures to the input data. The requirements for the second component were elicited by reviewing the properties of thirteen similarity measures found in literature, extracting their common features, and deriving requirements from their common features. The inspected similarity measures were used in a wide range of disciplines such as protein function prediction in biology, comparison of Web Service definitions in computer science and overlap calculation in graph theory, to name just a few application scenarios.

Before discussing different kinds of similarity measures, it is necessary to take the goal of their use into account. IS that make use of the similarity of task sequences can have different *target functions*. For example, in project controlling it is often relevant to assess the likeliness of success. This can be done by determining similar projects that have been completed already, taking their success as an indicator for the currently running project. In that case, the goal is to make a good estimation about project success. Another example with a different target function can be found in product recommendation engines where users are presented with similar products that overlap with their peer's preferences. In this case, the goal is to leverage the cross-selling potential. Different target functions have different definitions of when a similarity measure works well on a set of task sequences and when it does not. It is often appropriate to adapt a similarity measure to suit its intended support for a goal, using *supervised learning techniques* (Witten, Frank 2002), that use a learning set with labels to deduce results that can be generalized. The framework's similarity measure component, therefore, should fulfill the following requirement:

> The framework should have the capability to label a training set of task sequences with an indicator of its utility in relation to the target application's goal (*R.2a*)

There are a number of different ways to determine the similarity between two entities and also the *level of abstraction* of the relevant entities play a major role as also elaborated on in section 5.5.2.5. For this reason, the framework needs to be flexible enough to support each different way. In a first instance, an entity can be described by certain flat attributes, for example, the number of project members and the total budget may describe a project. In that case, one can compare the two entities according to their attribute values, where the comparison can be done with different algorithms depending on, for example, the data types or data ranges. This is why the framework should fulfill the following requirement:

> The framework should be able to support similarity measures that operate on input entities that are described by attribute-value pairs (*R.2b*)

In a second instance, an entity can have *structured components*, for example, a project is described by the tasks that have been performed during its execution. In that case, the two entities can be compared according to the overlap of the same constituting parts, i.e. the same tasks. The framework, therefore, should fulfill another requirement:

> The framework should be able to support similarity measures that operate on structured input entities by for example comparing the overlap of components (*R.2c*)

Additionally, the *constituting parts can themselves have attributes*, for example each task in a project can have a specific person that is responsible. Therefore, the comparison of entities can be based on constituting components, acknowledging the difference in attributes as well. Essentially, this is an extension of the requirement described before, where the constituting components were treated as flat structures and were compared for equality. The framework should fulfill this derived requirement:

> The framework should be able to support similarity measures that operate on structured input entities where each structured component is (additionally) described by attribute-value pairs (*R.2d*)

Finally, it is possible to take the *relationship between the constituting parts* into account. The relationship represents the temporal or logical order of the constituting parts and may also reflect interleaving of those activities. Addressing this fact, the framework should fulfill the following requirement:

> The framework should be able to support similarity measures that operate on the structure of its entities, i.e. that use structural properties of the input data for similarity determination (*R2.e)*

Additionally, many similarity measures use one or more *parameters to configure the computation* of similarity. The framework, therefore, should fulfill an additional requirement:

> The framework should offer the capacity to process parameters for each similarity measure that determine its behavior (*R.2f*)

Each similarity measure typically focuses on one or at least a small set of properties of the input object. However, it is possible that the desired notion of similarity is best reflected by a combination of different properties. In this case, the *simultaneous application of different similarity measures* is necessary. This poses another requirement to the framework:

> The framework should allow for a compounded calculation of similarity using different measures (*R.2g*)

As indicated above, the *structural properties* of task sequences can be used for the similarity determination. Yet, each task sequence itself has a linear structure by definition. To find out about the dependencies between activities, a model of possible sequences indicating their relationship is required. In many cases, explicit models of task sequences are not available because they are too expensive to create or because the task sequences are too flexible to render a model useful (van der Aalst et al. 2004a). Nevertheless, if the usage of structural properties is deemed necessary, there needs to be a way to at least *recover an implicit model* for the task sequences. It would have to be reconstructed from the IS logs and would then indicate the process "as it is lived". In terms of similarity determination, it can be used to *deduce structural properties* of an otherwise linear task sequence. The framework should fulfill another requirement:

> The framework should provide a possibility to create a (process) model using the task sequences that are available (*R.2h*)

As reasoned above, similarity measures only have a purpose with respect to their target application. Therefore, the framework should facilitate the selection and configuration of an appropriate measure. To find out about the utility of the selected measure and its configuration, the results have to be seen in the light of the application that they will be used for. The framework needs to be integrated into the target implementation or needs to be integrated into a suitable representation thereof to show its utility. Especially if supervised learning techniques are used for the selection and configuration of similarity measures, the *feedback of the application* about how well the similarity measure's results are suited for the intended use is pivotal. Therefore, the framework needs to fulfill a requirement that intentionally covers a broad spectrum of interpretations to encompass arbitrary similarity-based applications:

> The framework should support the integration of similarity-based applications or suitable representations thereof that consume the results of the similarity determination and give the framework feedback about the quality of the results (*R.3*)

The collection of the derived requirements is summarized in Table 6-1.

Id	Requirement
R.1a	Retrieve data stored in files and in databases
R.1b	Retrieve data from source that are structured, semi-structured or use structures in a proprietary format
R.1c	Handle semantically rich data sources and extract or amend semantically less rich data sources
R.1d	Be agnostic to how the data is captured from a technical point of view, but be able to amend data with implicit information
R.1e	Transform log information containing status snapshots into log information containing status changes
R.1f	Enable flexible and iterative enrichment of log data from multiple sources
R.2a	Label a training set of task sequences with an indicator of its utility in relation to the target application's goal
R.2b	Support similarity measures that operate on input entities that are described by attribute-value pairs
R.2c	Support similarity measures that operate on structured input entities
R.2d	Support similarity measures that operate on structured input entities where each structured component is (additionally) described by attribute-value pairs
R.2e	Support similarity measures that use structural properties of the input data for similarity determination
R.2f	Process parameters for each similarity measure that determine its behavior
R.2g	Support compounded calculation of similarity using different measures
R.2h	Support creation of a (process) model using the task sequences that are available
R.3	Support integration of similarity-based applications or representations thereof that consume the results of the similarity determination and give feedback about the quality of the results

Table 6-1 ***Summary of requirements for the framework***
Source: Own Illustration

6.2 The Framework's Architecture

The requirements discussed above informed the design of the framework. The logical structuring in three different components (see Figure 6-2) proved useful for the elicitation of requirements. For the implementation of the necessary functionality, it turned out that the second component can be split into three modules: one that supports *classifications* for supervised learning, one that can *mine a process model* from input logs for the support of structural similarity measures and one for the *application of similarity measures* itself. Therefore, the architecture features a modular design with five main modules (see Figure 6-3). The framework handles the flow of action by instantiating one or more plug-ins for each module and passing on the control subsequently. However, it is not mandatory to use all modules, i.e. classification of task sequences is only necessary when supervised learning should be supported and the creation of a task sequences’ model is only necessary if structural properties should be used in the similarity determination. The possible paths through the application’s modules are depicted in Figure 6-4.

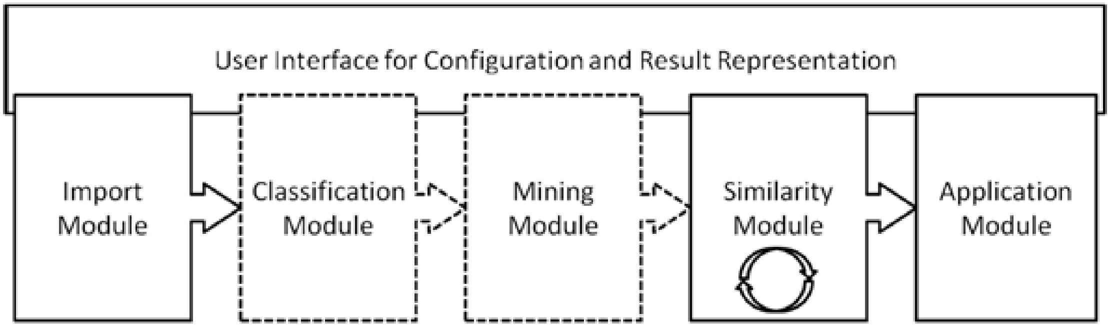


Figure 6-3 ***The framework's modules***
Source: (Schmidl et al. 2011b)

The framework was implemented using Java as programming language. The choice was motivated by the fact that Java is a very common programming language that many application engineers can work with, which allows plug-in creation for a larger audience. Wherever appropriate, existing applications and frameworks were integrated into the framework directly. This is true for parts of the import module that builds upon the ProMImport application and the process mining module ProM, that removed the effort of implementing process mining algorithms. If the functionality of the incorporated applications did not entirely suit the framework’s needs, they were extended to meet the set of requirements as elicited in the previous section. In this section, the framework’s five modules are detailed. The reasons for the design choices in each module are explained by reference to the requirements in section 6.1 that are also listed summarized in Table 6-1.

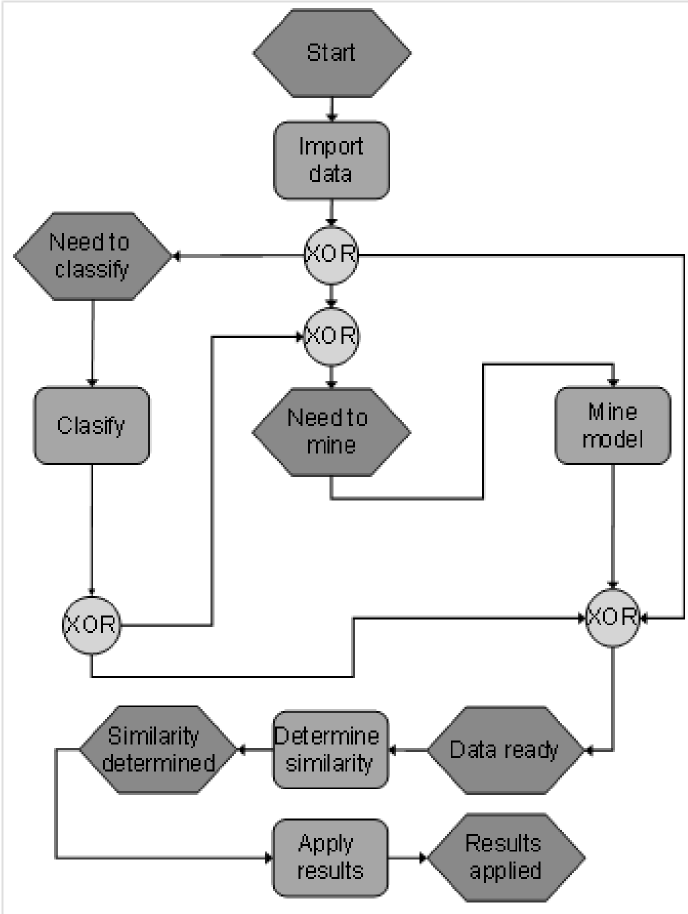


Figure 6-4 ***Possible usages of this thesis' framework for similarity determination***
Source: Own Illustration

Each plug-in of each module can be configured. Depending on the plug-in, there can be various options to fine-tune the behavior. For example, each similarity measure can indicate how the similarity of attributes should be determined, in case a task has attributes. All configuration options are entered making use of the framework's graphical user interface, where it is also possible to load and store configurations. A screenshot of the framework's interface that shows some of the configuration options for similarity measures can be seen in Figure 6-5.

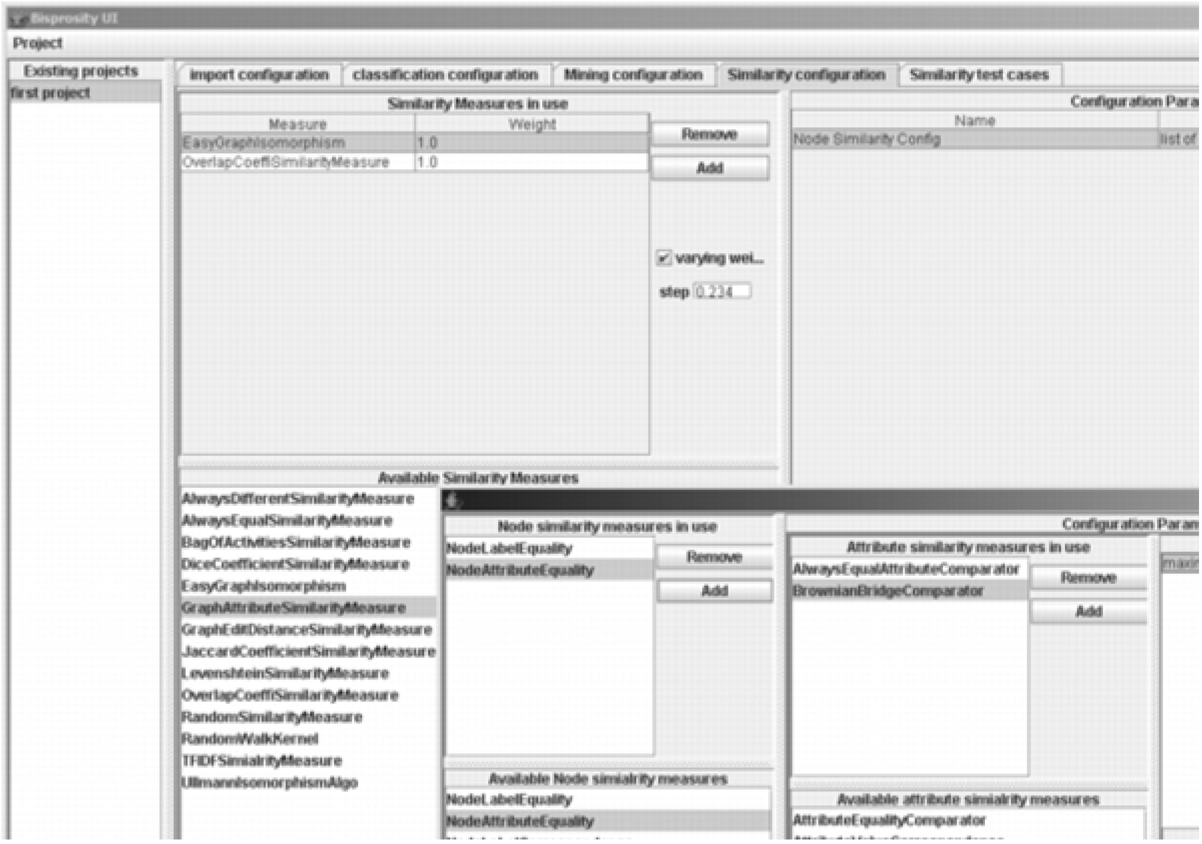


Figure 6-5 ***User Interface of the framework for the determination of similarity measures***
Source: Own Illustration

6.2.1 Importing Data into the Framework

The requirements analysis showed that task sequences are frequently stored in different locations, i.e. databases or files, and have different formats. Therefore, importing data into the framework for similarity determination cannot be done with one single import routine but rather must be done by offering an *interface* that supports as much flexibility as possible. A plug-in that implements this interface then offers the application-specific ability to extract task sequences.

Extracting information from logs is a common challenge. An existing tool, the ProMImport Framework (Gunther, van der Aalst 2006), served as a good basis for the import module, although its extraction mechanisms primarily focus on the support of process mining. For its use in the framework, ProMImport had to be extended. Specifically, the ability to store the extracted data before displaying it and the ability to connect the output of one import plug-in to the next import plug-in was added. In this way, requirements *R.1a* and *R.1b* are addressed by delegating the specifics of the data extraction to plug-ins while offering a generic interface.

The requirements analysis also shows that relevant information for task sequences could be spread across different logs that logically complement the information contained in each. The

ProMImport Framework had no support for aggregating information found in different kinds of sources for one combined import result. Therefore, the concept of *chaining importer plug-ins* has been implemented into the framework. While ProMImport would display the results of the import directly, this is not necessary for the framework developed in this thesis. Consequently, the results are not directly written to the output, but possibly augmented by the results of other, chained importers before the aggregated result are then finally passed on into the framework's chain of processing. Additionally, the result is stored into the file system. This dual approach was motivated after the first tests of this module had been performed. It turned out that the transformation into the internal format, thus the data importing, takes a considerable amount of time. As tests indicated, the transformation itself took the majority of the necessary time, while a simple read-in of data in the correct format was done comparably fast. Therefore, a plug-in was implemented, that makes correctly-formatted data available to the framework quite fast. To use this *plain reader* the import of other data sources has to be performed once before the result can be reused in later runs. This fulfills requirement *R.1f*, which states that information could be spread across different sources.

Offering the ability to connect arbitrary data sources also requires that a suitable data format is defined within the framework. As there are different alternatives that need to be investigated, a dedicated analysis of the alternatives and the final selection of a suitable data structure is necessary and is discussed in section 6.3.1. The result of the analysis and selection – the *MXML* format – is flexible enough to handle logs with varying granularity and is abstract enough to handle logs with varying contextual data. This fulfills requirements *R.1c* and *R.1d*, which state that the granularity of data can vary.

Additionally, the requirement analysis showed that there might be data formats storing events while others store statuses without the events leading to them. To cope with this challenge, section 6.3.2 describes the implementation of an additional tool to transform both formats into the MXML format. This fulfills requirement *R.1e*. In addition, the tool allows to assign attributes and their values can be calculated using basic arithmetic and string operations. This fulfills requirement *R.1f* that asked for the ability to amend the log data.

6.2.2 Classifying Task Sequences

The requirement analysis specifies that a label must be assigned indicating the *utility with respect to the desired target*, if supervised learning is to be used. In principle, two different ways to allow for labeling are possible: *automated or manual*. In one case, task sequences are labeled according to one or more rules that are created by a domain specialist. For example, the result of project executions labeled as good, mediocre or bad could be automatically determined using the budget-to-spending ratio of each project as basis for a rule. However, this automated approach has a significant disadvantage. If there was a rule available that perfectly labels this type of task sequences, then this rule is at the same time a perfectly suitable similarity measure and there would not be any need for using the framework in the first place. However, the more complex task sequences are, the less likely it is that a person knows according to which measures a task sequence should be evaluated. This person nevertheless may often be able to indicate the result *tacitly* without knowing how to derive this judgment formally. Therefore, the framework uses the other choice, namely the *manual labeling approach*. In this case, a person classifies a training set of the task sequences and stores the results in a csv file. This file is then used as an input source for the framework to automatically classify the training set of task sequence. As outlined in the requirements analysis, labeling is not always necessary and is implemented as an optional step in the framework. Having the labeling module fulfills requirement *R.2a*.

6.2.3 Generating a Model of the Task Sequences

The requirements analysis shows that many applications do not have an explicit model for the execution of task sequences, but its users may follow implicit models that, for example, stem from corporate rules or from the technically possible interaction offered by a graphical user interface. If the implicit model can be explicated in a potentially only approximated model, it can be used to extract structural properties of an otherwise linear task sequence. Inductively creating a model from instances contained in a log is the main concern of the *process mining* discipline (Agrawal et al. 1998). There are many algorithms available to mine a model from instances. The application of these algorithms is facilitated by the ProM framework (van Dongen et al. 2005; Verbeek et al. 2010) that has many of them integrated as plug-ins already.

Yet, ProM's abilities had to be adapted to fulfill all requirements. Since MXML will serve as internal data format, the input routines did not have to be altered– MXML is ProM's indigenous data format. On the model miner's output site, the result had to be adapted with considerably higher effort. The goal of the model miner lies in displaying a retrieved process model, which is why the only requirement for the output format is to be displayable. For the purpose of this thesis' framework, however, this does not suffice. Model mining is just one step in the framework's processing chain and, therefore, there was the need for the output to be in further processable format. Additionally, ProM returns its mined models typically as *event-process-chains (EPC)* (Keller et al. 1992) or as *petri-nets* (Peterson 1981). This makes sense as one of the important aspects in model mining is understanding implicit execution semantics. In the context of determining similarity, the execution semantics do not play a large role though, which allows for the simplification of petri-nets and EPCs into *simple graphs* that only consist of edges and vertices. This also creates the ability to use similarity measures that work with simple graphs. Furthermore, the algorithms used in the ProM framework are configurable to varying extend. Therefore, the configuration interface was adopted in this thesis' framework so that ProM's algorithms can be configured through the same interface as the remaining modules. In section 6.4 the research area of model mining is briefly introduced and also justification is provided why the simplification of petri-nets and EPC to simple graphs is adequate for the purpose of this thesis. Utilizing the existing ProM framework but adapting its output fulfills requirement *R.2h*.

6.2.4 Determining the Similarity between Task Sequences

The requirements analysis indicates that the framework must offer an interface for the creation of own similarity measures. Nevertheless, it is desirable to have a reasonable number of algorithms available in the framework to make it useful from the start. As the algorithms differ with respect to the parts of the task sequence they use for the computation, the framework should offer at least *one algorithm for each kind of similarity measure*. This guarantees that other similarity measures that operate on the same kind of input data can be integrated into the framework.

The framework has 13 similarity measures integrated in its initial phase, that can work with the general properties of a task sequence (requirement *R.2b*), the overlap of activities (requirement *R.2c*), taking into account the activities attributes if necessary (requirement *R.2d*) and also taking into account structural properties (requirement *R.2e*).

Within the framework, the main task of each similarity measure is to determine a *similarity matrix*, i.e. it must create a matrix with as many rows and columns as there are task sequences with each entry containing the degree of similarity between the respective combinations of task sequences. This very general representation of a similarity measure's result allows applications to extract the relevant information flexibly. A downside is that this way of storing a similarity measure's result is not the most efficient way considering computational cost and space requirements.

The requirements analysis also reveals that similarity measures frequently need configuration to a certain degree. The framework provides the means to properly configure a similarity measure. Each similarity measure plug-in is requested to publish its *necessary parameters* to the GUI component and can then process them as needed. This fulfills requirement *R.2f*.

Further, the requirement analysis shows that cases must be supported in which one similarity measure is not enough. A *combined result* of different similarity measures might fit the application's needs better than a single measure could. The framework supports this kind of configuration. Internally, each measure computes one similarity matrix. The matrices are then combined to yield one similarity matrix. The combination is done by also allowing for a weighting between the measures. This fulfills requirement *R.2g*.

As discussed before, depending on the abstractions level (see Figure 5-18), task sequences can be compared using attributes describing the sequence itself, or taking into account the single tasks contained in the sequence, or in a third way also considering the attributes of the tasks. In the first case the determination of similarity focuses on comparing basic data types that are used for instance description. The same holds for the highest granular view in which attributes are compared and again those are in the end comparisons of basic data types. This is why in the following some basic ways to define similarity of attributes are presented. Moreover, when taking the task level as relevant entity, the question of when two tasks should be considered equal becomes relevant, which is why it is addressed in the following as well.

6.2.4.1 Similarity of Tasks

As this thesis' focus lies on finding similarity between two individuals by inspecting their tasks, it is necessary to clarify in which ways tasks can generally be compared. In the following, the word task always refers to the computerized part of a task. The most obvious way to compare tasks, is to deem them equal if they have the *same name*, thus have the same "purpose". Comparing tasks this way is justified if the task's attributes are either of low importance or if tasks are usually executed with the same attribute values anyway. Another way to compare tasks is to define *correspondence classes*, thus to define classes, in which each contained task is equal to the other members of its class. This makes sense if there are activities that are semantically very closely related. One might for example want to consider the task "send order confirmation via e-mail" to be equal to the activity "send order confirmation via EDI[6]". This is of course highly application-specific. Tasks can also be compared by assessing the *similarity of their attributes' values*. To this end, one can compare each of the attributes that are available in both tasks. The returned values are averaged or combined by a weighting function to arrive at a final value of similarity. If this value is above a threshold, the tasks are deemed similar enough to be matched, otherwise not. In the next section, possible ways of determining similarity of attributes are described.

6.2.4.2 Similarity of Attributes

Attribute similarity always works on attribute values. If an attribute is only present in one task and or instance (when looking at only the first abstraction level) but not the other, similarity with respect to this attribute can simply not be evaluated. One might also consider two tasks or instances as different if they do not share the respective attribute. Three notions of similarity for attributes have been identified and implemented into the framework. The first and most basic notion is asking for *exact equality*, thus, two attribute values are either exactly equal, otherwise completely different. This is a suitable measure for categorical data, like e.g. names of customers that are the subject of a task. A related idea is to define *correspondence classes*, i.e. one defines classes of attribute values that are considered to be equal. This might be a

[6] EDI – Electronic Data Interchange is a standard for message transmission between different IS

good idea if e.g. a small number of customers are in the same category such as very important customers. If one is interested in the class more than in the specific customer, one could set up correspondence classes. However, when dealing with continuous data the previous two measures are not very helpful. Dealing with those attributes, the use of for example a *Brownian-bridge kernel* or a *Gaussian* kernel may prove valid. The first one determines a similarity value for the two attribute values by checking their distance and using an upper limit for this distance above which the similarity is deemed to be zero. The second one uses a Gaussian distribution with provided standard deviation as its basis.

6.2.5 Applying the Selected Similarity Measure in the Target Application

The framework is designed to support any application that builds upon the use of task sequence similarity. As it is unknown which kind of target application will make use of the framework, this creates the need to offer *flexible configuration mechanisms*. One way of configuring the similarity application is the use of *incrementable parameters*. These parameters are set up with a maximum, a minimum and an initial value, along with a step size. The application iteratively performs its task and the framework changes the incrementable parameter as indicated by the step size, until the upper or lower limit is reached. This functionality is indicated by the circle below the similarity component in Figure 6-11. In what ways the parameter influences the application's result determination mechanisms is not controlled by the framework itself, but by the application plug-in, while the framework performs the increment and manages the results reflected by the application plug-in. If more than one incrementable parameter is set for the application, the framework explores every possible combination.

Another possible need of future applications that utilize the framework might be to act differently, *depending on the results of the previous steps*, i.e. depending on the result of the import from the original source, the possibly extracted interaction model, or the result of the classification. Therefore, the framework offers the application an interface for interacting with the *intermediate results*[7]. The concept is as follows: Directly after the intermediate information is created, the application is asked to pre-process the intermediate information, presenting all intermediate results through an open interface. This happens before the determination of weight distribution and before using incrementable parameters. If the application uses this option, it can tell the framework how to pre-process the dataset and by this means adapt it to its needs. It is then provided with the pre-processed data instead of the plain intermediate results. This step is performed as long as the application indicates that it still wants to change the data. This implementation is generic enough to support arbitrary applications but offers enough functionality to still support the application engineer which fulfills requirement *R.3*.

[7] The imported task sequences, the potentially mined model, and the potentially created classification are considered as intermediate results.

6.2.6 Making the Framework More Flexible

The goal of the framework is not only to relieve the application architect of the task of finding a suitable similarity measure, but also to find a well-suited configuration of the similarity measure. The similarity application module supports this feature. When the usage of more than one measure is desired, the framework can be used to *determine the best combination in terms of weighting*. The user only needs to specify how fine-grained the search of the best solution should be by providing an increment value. This value is then used to exhaustively search the result space, which is done by iteratively using each weighting combination for the similarity measures. The combined measures' result is determined in the light of the application that builds upon them, which in turn informs the framework how well this combination is suited to its needs.

After each possible iteration that might stem from either the utilization of *incrementable parameters*, the application's use of *pre-processing calls*, or *determination of optimal weight distribution* between more than one similarity measure, the application returns its collected information to the framework. All information from each iteration is collected in one encompassing data structure. Effectively, the collected information reflects the respective performance of each possible combination. For this purpose, it uses a multi-dimensional array, where each dimension represents one incrementable parameter, the preprocessing if applicable, or the weight distribution if applicable, while the array's value represents the parameterization's performance with respect to the application's performance criteria.

To enable the user to visually explore the relationships, the user can select a graph that shows a two-dimensional projection of the resulting multi-dimensional array. The two dimensions of the graph can be determined without limitation.

6.3 Data Model and Transformation

As stated before, suitable similarity measures are anticipated to be helpful for many applications that deal with instances of task sequences. Since the framework will likely have to deal with *different sorts of input data* and especially *different formats* thereof, it is necessary to think about how to cope with this requirement. This section addresses this requirement. It defines *an internal data format*, onto which the source system's raw data will be mapped. The reason for using the internal format is motivated, different formats are investigated and a choice between them is made, after comparing the two most promising options. The two deeply investigated formats will be *Common Workflow Audit Data* and *MXML*. The later one will be chosen, as it is better suited for processes that are not necessarily the basis for strictly defined workflows. Then another challenging issue is addressed. Some systems store data states, but not the events that lead to these states. As task sequences are however a description of events that result in states, it is necessary to find a way to reconstruct those events. A transformation step is therefore introduced that is capable of translating data stored in a *state-centric* format into data stored in an *event-centric* format. The transformation is achieved by a self-defined XML application, that offers high configurability to allow other data, stored in a comparable way, to be worked with, also. The syntax and the semantics of this application will be detailed.

6.3.1 The Internal Data Format

The framework needs an internal data structure that can be used for intermediate storage. This format will be the only input format for the framework itself, other formats are supported by providing adapters, i.e. plug-ins of the import module, that transform the data as necessary. As the requirement analysis indicated, logs can have *varying expressiveness* concerning the granularity of logged activities and concerning each activity's context information. Therefore, the intermediate internal format must be chosen carefully.

Firstly, the internal data format must be designed in such a way that the transformation of source data into this internal format is *without loss of expressiveness* or limits the loss to a minimum. Secondly, the format should be chosen in a way that the tasks which will be performed on the data are supported as best as possible, i.e. the data should be *easily accessible* in the way the tasks need. Thirdly, the format should in the optimal case be a *standard format* that is backed by a standards body, or should be as close as possible to such a format. The last point is motivated by two facts. On the on hand and most obviously, using an (open) standard allows for better interoperability, i.e. the framework will be more likely to be accessible from other systems and therefore there is a better chance of getting data in the appropriate format. On the other hand, the standards bodies offer the possibility of information exchange for the involved parties, that leads to the fact that those standards are way more likely to take into account the necessary aspects of diverse source systems, are tested in real life applications and are less error-prone than a format that is developed by only one company, let alone only one person.

Data formats that are crafted like this can be found in the WFMC's specifications of the Common Workflow Audit Data (CWAD) (WFMC 1998) and in the MXML format (van Dongen et al. 2005). Both data formats are abstract enough to represent the contents of different application logs. Being tailored to the needs of a workflow system, the CWAD format has a considerable number of attributes that only apply in a workflow context. The MXML format on the other hand, abstracts from workflows and therefore allows a more straightforward transformation of arbitrary data. The following two sections will investigate both formats briefly which will inform the decision of which of the two to use as internal data format for the framework.

6.3.1.1 Common Workflow Audit Data (CWAD)

The following format was issued by the Workflow Management Coalition (WfMC). It is an organization consisting of more than 300 workflow vendors, users, analysts and university/ research groups. “The WfMC creates and contributes to process related standards, educates the market on related issues, and is the only standards organization that concentrates purely on process.” (WFMC).

Among other specifications that are to ensure interoperability, the WfMC issued a specification for the format of audit data. It is used to clarify which information needs to be recorded, in order to enable the reconstruction of what events where processed by a workflow system. The document (WFMC 1998) defines on an abstract level what must and can be part of the XML-coded data to be compliant with the CWAD specification. In the following the relevant parts of the specification will be pointed out. The fundamental unit of information is considered to be an event. An event is either a request or a response or a status change within a process instance. Each logged event needs a prefix that contains the information as shown in Table 6-2.

Attribute name	Interpretation
Initial process instance ID	Unique ID of root process instance
Current process instance ID	Unique ID of current process instance
Activity instance ID	Unique ID of current activity instance
Process state	Current state of the process instance
Event Code	Event code value
Domain ID	ID of domain of corresponding user
Node ID	ID of node of corresponding user
User ID	ID of responsible user (person or entity)
Role ID	Role of involved user
Timestamp	Time at which the event was recorded
Information ID	Type of information

Table 6-2 Attributes in the CWAD standard for an audit log entry
Source: (WFMC 1998)

One distinguishes between initial process and current process whenever it occurs that one process instance calls another process instance. It is not mandatory to specify the *Activity ID*. The *Process State* indicates if the process instances is currently active, scheduled or on hold. *Event codes* tell what action has been performed, e.g. an activity instance has been terminated. It suffices to specify either *User ID* or *Role ID*. Together with the *Domain ID* and the *Node ID* the entity who issued the event can be exactly identified. The timestamp is used to be able to chronologically order events possibly among different systems. The *Information ID,* is used to tell if the information contained in the remaining part of the log entry is either private, i.e. vendor-specific, or WfMC conform, thus the information must be logged for this event type. Whenever process instance or activity instance data is updated or added, the CWAD suffix becomes obligatory. It consists of the attributes as list in Table 6-3.

Attribute name	Interpretation
Account Code	Accounting code used for this work item
Extension number	Number of extensions in the suffix
Extension type	Type of extension
Extension length	Total length of extension values
Extension codepage	Codepage used to record extension data content
Extension content	Content as defined by type and length

Table 6-3 ***Suffix attributes in the CWAD standard for an audit log entry***
Source: (WFMC 1998)

The account code is optional. If there are no extensions, the number of extensions is set to 0. Depending on the event, some additional attributes must be specified. For example, if a new process instance is started, the *Process Definition ID* of the instance's process must be specified. Another example is the update of a process instance's state, in which case the old and the new state must be supplied. Since the specifics do not contribute much to the understanding, the interested reader is referred to the original specification (WFMC 1998) for more information.

6.3.1.2 MXML Format for Process Mining

Another proposal for a common audit log format, called the MXML format, was proposed in (van Dongen et al. 2005). It mainly aims at supporting the task of *process mining* (see section 6.4) and corresponds to the specification detailed above in some ways. Figure 6-6 shows its UML diagram representation. It is used to map proprietary workflow system's concepts and terminology onto the MXML format's concepts and terminologies. Although WFMS are the most important source systems, they are not exclusively aimed at, but all kinds of PAIS are addressed. Interpreted broadly, even a code version management system can be seen as part of this class of systems.

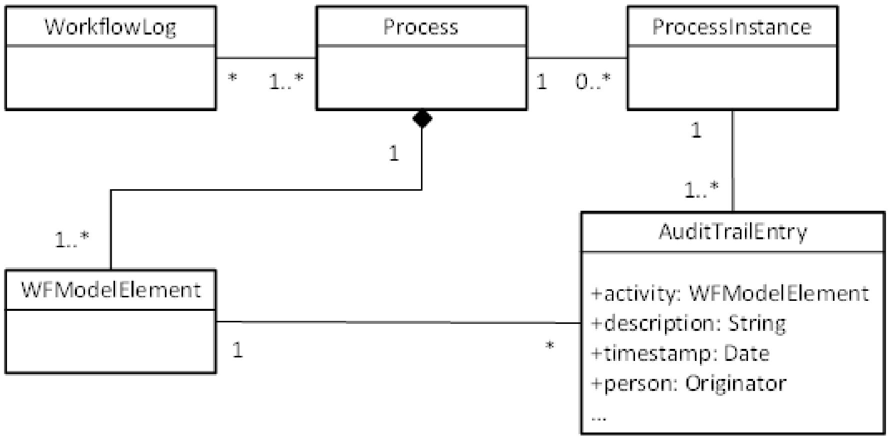


Figure 6-6 ***Meta model used to map onto the MXML format***
Source: Adapted from (van Dongen et al. 2005)

As can be seen in Figure 6-6, a log contains at least one *event* of a certain process. A *process* groups different *process instances* that have the same goal. A process could for example be the "software order process", while process instances would then be orders of e.g. an office application or an operating system. A process consists of several *activities* (in the diagram depicted as *WFModelElements*; these terms are used synonymously), while a process instance has one or more *audit trail entries*, which correspond to *events*. Each audit trail entry is associated with a task, thus each event induces a state change of a task. State transitions of tasks are conducted according to the state diagram as seen in Figure 6-7.

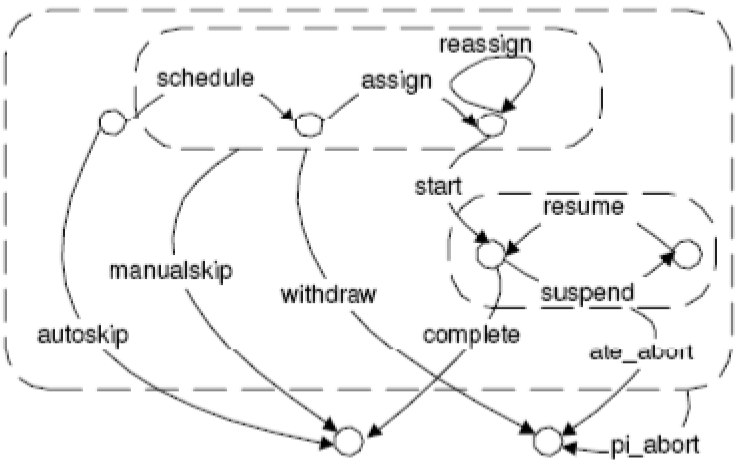


Figure 6-7 ***State diagram of MXML tasks***
Source: (van Dongen et al. 2005)

Dongen et al. (2005) created the state diagram after evaluating the possible states of various PAIS. Usually, different systems support different subsets of the state transitions. In the minimal case, one can also only consider the state *finished*. In order to include arbitrary data on every conceptual level and to allow for an easy extendibility, the MXML data schema provides for an optional data tag. The complete XML Schema of the MXML format can be seen in Figure 6-8.

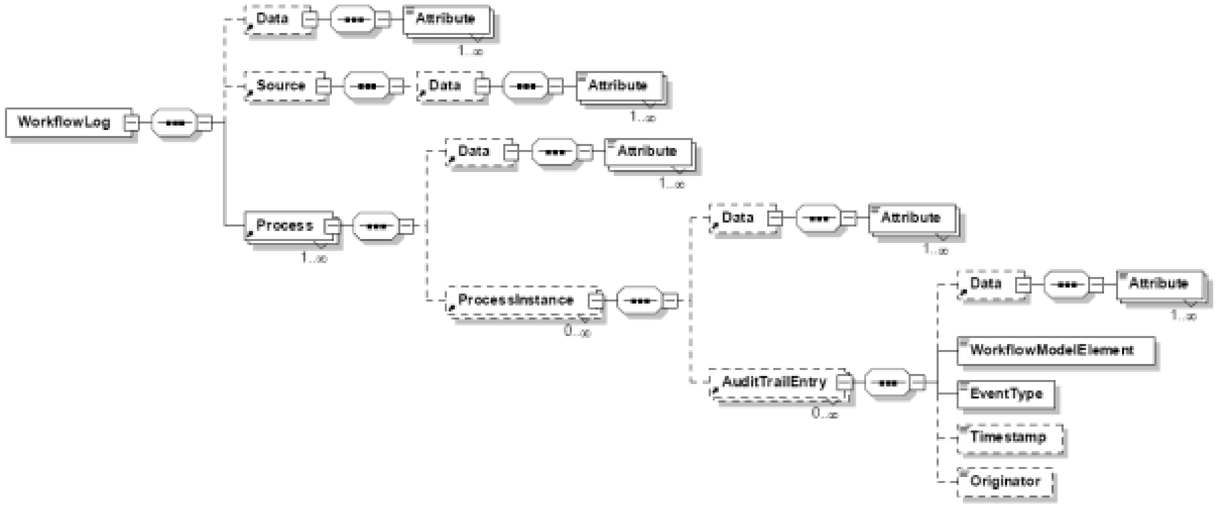


Figure 6-8 ***XSD Schema of the MXML format***
Source: Adapted from (van Dongen et al. 2005)

6.3.1.3 Comparison of the MXML and the CWAD Format

When comparing WfMC's CWAD specification for audit logs and the proposed MXML format, one can see some correspondences and some differences. At a first glance, the two proposals differ in their structure. The MXML format is explicitly *hierarchical*, thus an event belongs to its containing process instances, which in turn belongs to its containing process. The CWAD format on the other hand relies on a *reference paradigm*, thus each event includes the IDs of its activity, user and process, thus building on referential dependencies, similar to a relational database design. However, this difference is in data structure only, while the expressiveness of both formats is the same. The one format can easily be transformed into the other and vice versa.

At the top level, the MXML format expects a workflow log, which has no explicit correspondence in the CWAD format, but this is rather a formality than a conceptual difference. MXML expects a workflow log to consist of different processes. This is consistent with the CWAD format that also expects process identifiers for each process instance. The CWAD format explicitly allows call hierarchies of process instances, which has no correspondence in the MXML format that does not consider processes instances calling other process instances at all. Both formats consider process instances to contain activities, the CWAD format by using IDs, the MXML format by hierarchical tag inclusion.

A significant difference lies in the fact that the CWAD format has another logical unit, the *work item*. An activity consists of work items, which are assigned to users of the workflow system. MXML considers *tasks* as the fundamental unit. Although they need not be considered atomic, e.g. by using the task scheduling schema appropriately, they have no resemblance to work items.

On an attribute level, there are some more differences. The MXML format allows for an *originator tag*, that contains the person or entity that triggered the event. The CWAD format on the other hand has a richer means of telling who triggered an event by asking for a *domain*, a *node*, and either a *user or a role ID*. Still another difference is the two formats' scheduling schemas. While the CWAD format relies on a process instance-wide state attribute and activity state attributes, the MXML format uses a scheduling schema for activities only. Furthermore, events in the CWAD format are determined by an identifier and represent events like "create process instance" or "change process instance state", thus *predefined events*, the MXML format only knows events that are one of the *scheduling transitions combined with a task*, like "start task1" or "complete task1". A last difference lies in the format of additional data that might be carried in the log entry for an event. In the CWAD format, *arbitrary data* can be stored[8], with some *meta-information* about the additional data, like what kind of data it is. The MXML format only knows the *data tag*, which may contain arbitrary, text-coded data, but without specific meta-information about it, except the meta-information that could be conveyed through XML tags.

Both formats have in common that they store a *timestamp* for the logged event. In sum, one can conclude that both formats are closely related. While they follow different storage paradigms, namely hierarchical, XML storage and referential attribute constructs, they both resemble the same basic structure. However, the CWAD format often provides richer means for specifying e.g. a user or a related process instance. While the CWAD format expects a given process model, the MXML format is designed to only ask for a process model name and, by utilizing process mining, find a model for the given log data. In essence, the focus of the two formats is different and naturally follows the goal of the respective applications. MXML aims at supporting the process mining task as best as possible, while CWAD tries to best support the needs of a productive workflow management system. This distinction is supported by the fact that MXML does not know CWAD's basic unit of structure, the work item, which is essential for a productive system, yet can be abstracted from when mining a model.

This thesis' framework will rely on the MXML format as internal data format. The reason for this choice is twofold. First, since the main goal of supporting the mediation of knowledge workers does not necessarily build upon a workflow management system, a mapping onto the CWAD data format might require filling many attributes with dummy values. Most potential systems only differentiate between users but might not know the concept of domains, nodes

[8] As long as the data type builds upon CWAD's basic types that are almost identical to those known from common high level programming languages

or roles, for example. In addition, many PAIS could potentially be used for this thesis' purpose. In general, there is the tendency that the more knowledge-intensive a process is, the less stringent a process model should or can be, often leading to the inability to design a model at all. Effectively, it is often the case that there are no models in this thesis' field of application. The CWAD format, however, expects a (process) model. As indicated in the outline of the framework's requirements and its architecture, it is sometimes necessary to recover a model by using process mining techniques. To use these techniques, the MXML format is best suited, especially since there are existing tools that build upon this format, which allow for a high degree of freedom.

6.3.2 The Data Transformation Approach

This section addresses the issue of how to transform data that follows another paradigm, that shall henceforth be called the *state-driven format*, into data that follows the *event-driven format* that is used in the MXML data format and, therefore, also in the framework. In the following, the necessity of such a transformation is motivated. Then the approach of achieving such a transformation with the help of a self-defined XML application is detailed. The syntax and semantics of the transformation are explained and the hierarchical design is illustrated. The transformation is necessary due to the fact that activities, that are part of events, are the essential part of process-oriented applications and some input formats simply do not contain events as such in their format. The requirements of the internal format and the anticipated input formats lead to an XML application that uses XPath in the back-end and is highly configurable.

6.3.2.1 Necessity of a Transformation

The MXML format expects the data to represent a sequence of events, i.e. in an *event-driven format*. Thus, each event, its name and the effect it has on the data is stored. An event could for example be "increase cost by 5,000 €" that increases the cost attribute. This is illustrated in Figure 6-9 on the left side. Yet, there are also systems that do not store events but a data set, or *image before and after* some event, thus essentially states, hence it is referred to as *state-driven format*. Sticking to the example before there might be a state S_1 in which the cost of a project is 5,000 € and another state S_2 that was created after some unnamed and unknown interaction with the system in which the cost of the project is then 10,000 €. This is illustrated in Figure 6-9 on the right side.

The two formats have a complementary view on the process execution. The first one, the event-driven format, describes what action was performed on the data. The state-driven data format on the other hand does not store the event itself, thus there is no information about what was done with the data, but only about the data state before the change and after the change, and therefore there is only information about how the data changed. There is a special form of the state-driven format that does not exactly store the data states before and after an interaction but only the initial state and then deltas of the new states and the previous states. This is similar to the event-driven format. Nevertheless, there is still no way to know what was done, only the effects can be seen. Clearly, one can conclude what happened by looking at the initial state, i.e. before the interaction with the system and the final state. In the example above a human being, can easily conclude, that in the second state the project cost is 5,000 € higher, thus the interaction must have been “increase cost”. However, without further information this is not possible for a non-human entity and this is what makes the use of the two formats a challenge. Process executions rely on states and activities, yet, a state-driven format only contains the states. It is therefore necessary to tell the application how a certain change looks like, i.e. what pattern must be present to conclude a specific activity.

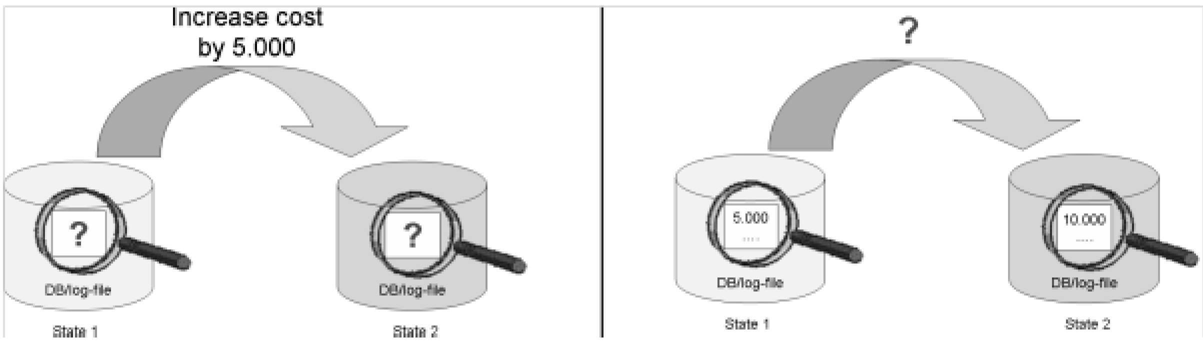


Figure 6-9 ***Data changes represented in the event-driven paradigm (left) and the state-driven paradigm (right)***
Source: Own Illustration

This problem has been addressed formally in theoretical computer science. A formalism called *situation calculus* has been developed in which one tries to use a logic framework to reason about situations and actions that transform one situation into another (McCarthy 1963, Reiter 2001). Although the formal side of this problem is highly interesting and still part of ongoing research, it does not lead to an implementation in a straight forward manner. Therefore, it is necessary to create an approach that offers a solution to this problem in a pragmatic way that will be described in the following. It allows to specify how to *transform* a state-driven to an event-driven data format in a principled way. This transformation is described in detail in the next section.

6.3.2.2 The Transformation

Since the input data and its format depends on the source system that creates the log data, it is not possible to create one transformation for all possibilities. In the following, a solution is described that is able to transform the general class of semi-structured input data stored in XML format. This is reasonable, as a fair number of applications store their logs in XML format. The principles that is described in the following also apply to other input formats, however the implementation would have to be different.

The transformation should be as flexible as possible. This motivated the creation of a XML-based configuration language that can be used to perform the transformation while allowing for a wide range of configuration options. The reason to choose an XML application for the configuration of transformations over key-value modeled or binary coded configuration specifications lies in its wide adoption. XML has become the standard format whenever interoperability between applications is important. XPath (W3C 1999) will be used as a selection language for the XML documents. Offering a rich variety of navigational steps through a document, it allows to select arbitrary structures, such as a node, an attribute, an attribute value or sets of the previously mentioned. Since XML is a quite established standard for data modeling, it is expected that many systems that are used in conjunction with the framework can make use of the transformation provided here. In addition, the output of those systems can be selected easily using XPath expressions.

A transformation configuration consists of a list of *mapEntry* elements. Each mapEntry includes a condition expressed in XPath. This expression is checked against the input source and for each match within the input source, an output entry is generated in the output of the transformation. There are two kinds of transformation definitions. The first one details how to map data changes to activities in a straight forward manner. For example it could state, whenever the input source contains a node *state1* with attribute *budget* and a node *state2* also with node *budget*, then insert into the output a node *budgetChange* with the attribute *difference*. The nodes *state1*, *state2*, and the attributes *budget* are searched for using XPath expressions. The content of the new attribute *difference* is computed during runtime. The application offers a broad set of fundamental operations on different data types. The second type of transformations depends on the first kind of transformations and uses their functionality by means of inclusion. Second type transformations can also use other second type transformations, effectively allowing to create a hierarchy of transformations.

After one-time configuration, the transformation application automatically selects defined portions of the source data that contains sequences of states and transforms them into event sequences during data import. The events are created using activity names that are configured before importing the data. By using XPath as selection mechanism, it is granted that the configuration can select arbitrary pieces of information from the input source. By offering diverse fundamental operations on basic data types, the application further allows to transform the input data in many ways. Finally, offering the ability to use inclusion mechanisms the transformation language can be used more efficiently. For details of the configuration of transformations, the interested reader is referred to the Appendix A.

6.4 Model Mining

The requirement analysis showed that there are similarity measures operating on *structural properties* of the input data, i.e. execution structure (requirement R.2e). However, the input data itself consists of *execution logs* that contain sequences of events, each belonging to a, possibly implicit, process instance. Thus, the events do not reflect any structure. In order to archive a structured representation the event logs need to be matched against a model, as shown in Figure 6-10.

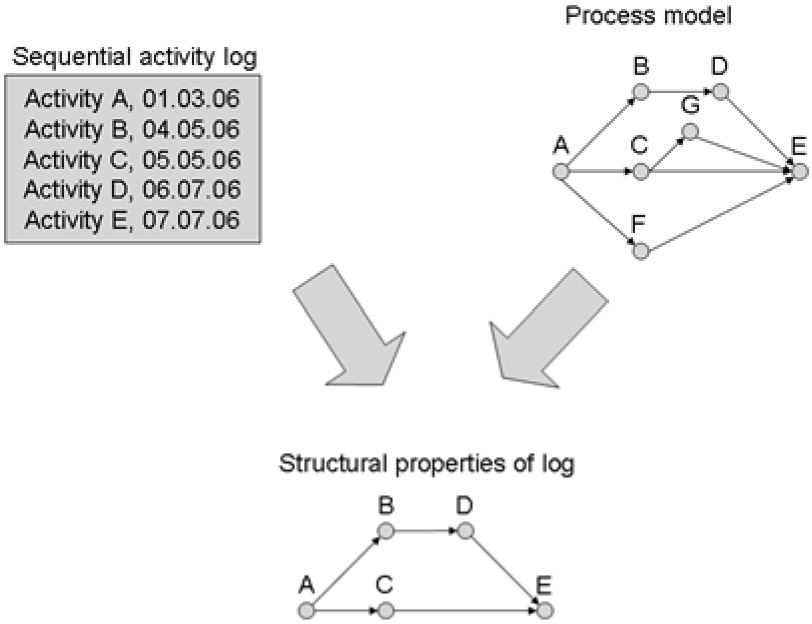


Figure 6-10 ***Derivation of structural properties of an interaction using a process model***
Source: Own Illustration

However, many operational systems, that could be used for determining task context overlap, do not have an explicit process model, but its users follow implicit process models that stem from e.g. corporate rules or from the possible interaction offered by a graphical user interface. The challenge of reconstructing process models from log files is investigated in *process mining research*.

6.4.1 Short Introduction to Process Mining

The general idea of process mining is to generate a suitable process model using system *interaction logs* from a PAIS. The interaction logs are usually in the form of process instance audit trails. Those audit trails can be seen as a process instance’s footprints and can be anything from workflow management system logs to web server logs or login data of any proprietary system that resembles a process in some form. Independent of whether PAIS use an explicit (strict) model or not, they usually log process relevant actions. Those logs can be used to reconstruct traces of system interaction, which can in turn be used to generate a model out of them that *reflects the actual system usage*. The first approach to process mining has been proposed by Agrawal et al. (1998), but research in this field has intensified in recent years. A number of algorithms have been proposed for this task. Process mining can be used for (re-)discovering process models from audit trail logs on the one hand and for checking on the conformance of modeled and lived process. There is suitable tool support for this task to be found in the ProM framework that offers mining, transformation and analysis tools, of which only the first will be used for this thesis' purpose. Transforming ProM's output – Wf-nets – into graphs is necessary for the setting of this thesis, but can be done by simply "cutting out" the necessary structure as illustrated in Figure 6-10. This is an adequate process as will be shown.

6.4.1.1 Use Cases

Generating a model from audit logs is helpful in diverse use cases. For example, it is often the case that processes in a company should adhere to a descriptive or prescriptive process reference model. Such a model specifies how people and organizations are expected to work. By comparing the prescriptive process model with the discovered model, discrepancies between both can be detected and used to improve the process. Process mining could also be used to compare how different departments use reference processes.

Hence, performing process mining can be used for obtaining two different goals. First, it is possible to generate a model for a system, that previously did not have an explicit one. Second, it is possible to compare the modeled, i.e. anticipated and the actually performed interaction with the system, which is called conformance analysis. A derivation of the second field of application, is the detection of typical interaction patterns, that can in turn be used to learn about the process in general or especially its performance. Another possible use is for the detection of frequent patterns in processes, to support a continuous process improvement as for example described in (Ngu et al. 2005). Process mining methods can also be used for non-business aspects. In the computer security domain, intrusion detection and auditing can be supported (van der Aalst, de Medeiros 2005b), while in the web mining domain analyzing usage patterns is a possible application for process mining. With respect to the needs of the similarity measure determination framework, process mining serves the purpose to extract a structure from process instances for structural measures.

6.4.1.2 Tool Support

There are a few research groups that dedicate their research effort to the field of process mining, of which one group around van der Aalst has issued an open-source framework for process mining support known as *ProM framework* (van Dongen et al. 2005). It facilitates the implementation of new process mining algorithms. In addition, it offers the possibility to mine models using already implemented mining algorithms, convert them into different formats like petri-nets or EPC nets and perform checks on them, like e.g. checking on the conformance of the model with respect to the linear temporal logic formalism.

Because interfacing with the ProM framework opens the door for the use of diverse mining algorithms, ProM was chosen as the heart of the framework's mining component. However, only the framework's ability to mine models will be exploited, while the diverse other features are not necessary. ProM is constantly being extended. At the time of implementation of this thesis' framework, there were six different process mining algorithms offered through ProM and therefore also available through this thesis' framework. All of them have been used to find their relative advantages and disadvantages. Nonetheless, as process mining itself is outside the scope of this thesis and only its possibilities are relevant in this context, the interested reader is referred to the original sources where each of the algorithms is explained in detail: the α-algorithm (van der Aalst et al. 2004c; van der Aalst et al. 2003; van der Aalst, de Medeiros 2005a), the α^{+}-algorithm (de Medeiros et al. 2004; de Medeiros et al. 2005b) the β-algorithm (aka. Tsinghua-algorithm) (Wen et al. 2004), the genetic mining algorithm (de Medeiros et al. 2005a; de Medeiros et al. 2005c,), the multi-phase mining algorithm (van der Aalst et al. 2004b), and the hierarchical mining algorithm (Greco et al. 2004; Greco et al. 2005).

6.4.2 Transformation of Execution models into Graphs

This thesis' framework represents execution models internally as *simple graphs*. A simple graph is considered to be a set of nodes that are connected by arcs. It does not, in contrast to e.g. petri-nets, have execution semantics. This is certainly a simple way of representing models. In addition, it allows to use similarity measures that operate on graph structures – a characteristic of some of the algorithms that were adopted from other research disciplines and that were integrated into the framework. However, most process mining algorithms that come with the ProM framework return their results in the form of petri-nets or event-process-chains. For the purpose of process mining in its original attempt to retrieve models of business processes this certainly makes sense. Retrieving, modeling, and displaying execution semantics is an essential part of this task. Yet, in the context of this thesis, the execution semantics is not relevant. Therefore, it is necessary to think about *how to transform petri-nets and event-process-chains into simple graphs*. In addition, the consequences of using "only" a simple graph need to be investigated.

The transformation is done as follows (see Figure 6-11). Whenever the ProM framework returns a petri-net[9], the framework needs to transform it by creating one node for each of the

[9] In the context of process mining and also similarity determination event-process-chains can be transformed into petri-nets without loss of information, which certainly is not true in general

petri-net's transitions (blocks in the figure). Those nodes are then connected to one another by inspecting which transitions are connected in the petri-net, where the term connected is interpreted as follows. Two transitions are connected if there is exactly one place (circles in the figure) in between them. If the model has explicit routing nodes (not shown in the figure; would be *XOR*, *OR* or *AND*, with the obvious semantics), then two transitions are connected if there is a sequence of zero or more routing nodes and one place in between them, but no other transition. Additionally, the resulting graph is extended with explicit *Start* and *End* nodes that are implicit in petri-nets where tokens indicate start and stop situations. Using this definition and this way of transforming the petri-net, it does not matter if one uses explicit routing nodes or implicit routing by means of petri-net firing semantics. Both will be transformed into the same graph. An example of implicit routing can be found in the petri-net on the left part of Figure 6-11. In that example, *Task 3* must always be executed in accordance to petri-net firing semantics.

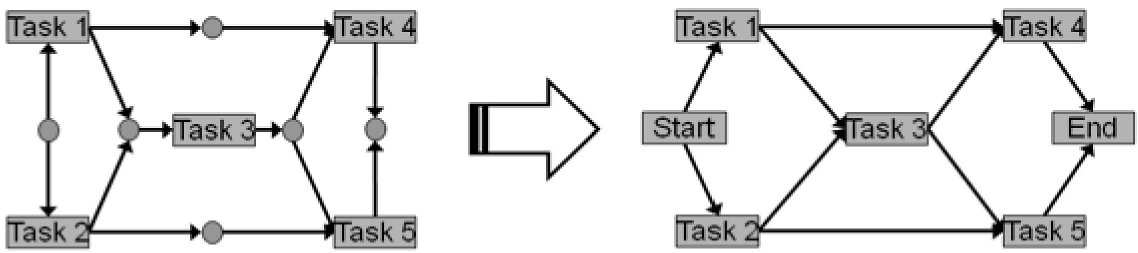


Figure 6-11 ***Transformation of a Petri-net to a graph***
Source: (Schmidl et al. 2011b)

The transformation approach is adequate. The argument is as follows. Since essentially places are removed from the petri-net, along with the routing nodes, if there are any, only the connection between transitions are left that are now considered nodes. This obviously explains the pure structural features of the transformation. It is, however, not obvious anymore if those connections are to represent exclusive execution or parallel execution. It is no longer known which transitions are prerequisites of certain nodes and, therefore, explicit information about the execution semantics is lost. Referring to the example in Figure 6-11, the petri-net forces the execution of *Task 3* before the execution of *Task 5* or *Task 4* respectively. In the resulting graph, a traversal from the start node to *Task 2* and then *Task 5* before finishing the path in the end node is possible. Apparently, during the transformation this difference in execution semantics is introduced. However, this is not a problem for the intended use. The transformed model is solely used *to identify the path that was already taken* by a process instance so far. Therefore, there is no need to know which connections previously modeled exclusive or parallel execution. The connections that are used by the instance were obviously parallel, the ones that have not been used but could have been according to the graph are of no interest, but obviously have represented OR/XOR split alternatives. It is also not a problem that there is no information about which nodes are prerequisites of another node, since the further proceeding of the process instance is of no interest – it is a purely *retrospective recreation of structural features*.

6.5 Summary

The approach of supporting knowledge transfer processes by building upon similarity of task contexts can be adopted more easily, when any effort is kept as low as possible. One substantial effort is defining the “correct” notion of what is similar. Having identified the need to back the approach by appropriate tool support, this chapter addressed the possibility to reduce the effort of finding suitable similarity measures for this purpose. It introduced a framework that allows to *(semi-)automatically determine suitable similarity measures*.

As a start requirements for such a framework were elicited by decomposing the overall goal into the sub-units *acquiring log data*, *application of similarity measures* and finally *checking the fitness* of the result. For the first sub-unit, data formats of contemporary systems were analyzed to deduce a set of six requirements for the proper acquisition of log data. Taking into account the properties of similarity measures and applied algorithms from other disciplines that use the concept of similarity in a next step eight requirements for the framework’s support of similarity measures were defined. The last sub-unit, checking fitness, is very context-dependent, which is why no requirements were elicited besides the one that arbitrary applications should be able to use the framework for their purpose.

Building upon the defined requirements this thesis’ framework for similarity measure determination is suggested. It consists of five modules: an *import module*, a *classification module*, a *mining module*, a *similarity module*, and an *application module*. Not all of the modules must be used; the only mandatory ones are the import, the similarity, and the application module. The functionality of each module was then described pointing out how it addresses the requirements elicited before. The overall architecture is *modular* in its design in that each of the five modules has clear interfaces and allows the integration of new plug-ins easily. Nevertheless, different plug-ins were integrated into the framework to make it directly useable. Altogether two import plug-ins, one classification plug-in, six mining plug-ins, twelve similarity plug-ins, and one application plug-in were part of the delivery version of the framework.

Classification is done using a list that indicates the utility of some instances with respect to the application. It is generated outside of the framework, for example manually. It serves the purpose of providing *training instances for supervised learning* approaches. Mining refers to the *retrieval of execution models* that are inferred from *interaction instances* from the log. This field of application and research is known as *process mining* and there is already adequate tool support. A well-known framework from this discipline, the ProM framework was included into the framework’s mining module. *Similarity determination* is supported for various kinds of algorithms. The framework allows to operate on execution instance meta data, the tasks of the instances, their attributes, or structural properties thereof. The usage of similarity results in a *target application* that is supported through a clean interface and different kinds of measures to *increase flexibility* and *increase the degree of automation*. The application can iteratively rely on *different weight distributions* of sets of similarity measures, can use *automatically incremented parameters* for all modules and can use the ability *to preprocess intermediate data* such as the retrieved process model, classification results or data import results.

The framework relies on a data format common in the domain of process mining. It was selected after carefully comparing the features with another common data format from the workflow support field. To address the elicited requirement, that two different kinds of data storage formats must be supported – *event-driven format* and *state-driven format*, first their principle difference was motivated. The event-driven format indicates what has happened, while the state-driven format indicates the result but not what has happened itself. The chosen data format MXML expects the data to be in event-driven format. The event-driven and state-driven formats can be transformed into each other, as verified by describing a *transformation language*. It is an XML dialect and uses XPath to select the relevant input sections, offers fundamental operations on various data types and creates output events in MXML format that are then usable by the framework’s internal mechanisms.

Furthermore, the requirement to support similarity measures that operate on *structural properties* has been addressed. Since it is, however, often the case that source system do not rely or have explicit process models, the possibilities offered by process mining are used to *deduce an implicit model*. Having this deduced model it is possible to “cut out” the structure of each interaction sequence from it. Another challenge lies in *transforming the data format* into a format as expected by the similarity measures: simple graphs. The process mining algorithms however retuned event-process chains or petri-nets. A suggested transformation approach was described and its adequacy was proven.

The framework fulfills all requirements that were elicited. It offers *adequate support* for one technically challenging aspect of similarity-based support for knowledge exchange, the *determination of suitable similarity measures* for the intended scenario.

7 Evaluation of this thesis' KM Approach

Evaluation of the artifacts that were created in a design process is a fundamental aspect in any design-oriented research (Österle et al. 2010, p. 4). Hevner et al. note that the " […] utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods" (Hevner et al. 2004, p. 85). The goal of this thesis was to design an artifact that would not face the challenges that are inherent in contemporary approaches for distributed, inter-personal knowledge exchange. For this purpose a KM approach was developed that builds upon a new way by *integrating two existing approaches*, namely knowledge networks and ERS, as well as a *matching mechanism for individuals* that operates on the *overlap of IT-supported tasks*. The evaluation of the resulting artifact may be done in different ways. A selection of methods compiled by Hevner (2004) is summarized in Table 7-1.

Category	Method
Observational	Case Study: Study artifact in depth in business environment
	Field Study: Monitor use of artifact in multiple projects
Analytical	Static Analysis: Examine structure of artifact for static qualities (e.g., complexity)
	Architecture Analysis: Study fit of artifact into technical IS architecture
	Optimization: Demonstrate inherent optimal properties of artifact or provide optimality bounds on artifact behavior
	Dynamic Analysis: Study artifact in use for dynamic qualities (e.g., performance)
Experimental	Controlled Experiment: Study artifact in controlled environment for qualities (e.g., usability)
	Simulation . Execute artifact with artificial data
Testing	Functional (Black Box) Testing: Execute artifact interfaces to discover failures and identify defects
	Structural (White Box) Testing: Perform coverage testing of some metric (e.g., execution paths) in the artifact implementation
Descriptive	Informed Argument: Use information from the knowledge base (e.g., relevant research) to build a convincing argument for the artifact's utility
	Scenarios: Construct detailed scenarios around the artifact to demonstrate its utility

Table 7-1 ***Design Evaluation Methods***
Source: (Hevner et al. 2004, p. 86)

Some of the evaluation methods are applicable for this thesis' artifact and research setting. Denzin (1978) explained that in qualitative research, *data triangulation* and *method triangulation* helps to achieve more elaborate results. While this thesis did not follow a qualitative research pattern, data and method triangulation can still help to increase the insights of the evaluation. Therefore, this thesis' central artifact – the KM approach – is evaluated using *multiple* *methods*. The evaluation starts with a *descriptive method* using *informed argument* to reason that this thesis' approach has advantages similar to contemporary methods but not their disadvantages. Additionally, an *analytical evaluation* – an *architecture analysis* – is performed to evaluate the ability to integrate an implementation of the approach in a large IT landscape. This also serves as a feasibility check for real life implementations and touches on performance aspects. Finally, next to the descriptive evaluation of suitability and the evaluation of technical implementability, the real-life usage of the approach is investigated. For this purpose two *observational evaluations* using two *case studies* with prototypical *implementations of the approach* were performed. The purpose of the case studies is to illustrate the possible benefits in a real-life setting. Additionally, the case studies are used to find out more about suitable and less suitable *environmental conditions* for the KM approach. This is necessary as not only the artifact influences its environment but also the environment inherently influences the effects of the crafted artifact. March and Smith stress that the understanding of the environment is essential by stating that " [...] significant difficulties in design science result from the fact that artifact performance is related to the environment in which it operates. Incomplete understanding of that environment can result in inappropriately designed artifacts or artifacts that result in undesirable side-effects. A critical challenge in building an artifact is anticipating the potential side-effects of its use, and insuring that unwanted side-effects are avoided." (March, Smith 1995, p. 254)

7.1 Descriptive Evaluation

Rigor in evaluating artifacts asks for an *evaluation against the goals* that were defined in the beginning of the research endeavor (Becker 2010). Becker notes additionally, that the benefit of the artifact for supplier or user should be described as artifacts are not designed for their mere existence but to *generate beneficial circumstances for its stakeholders* such as increased effectivity, efficiency, less defects, more transparency or shorter cycle times (Becker 2010).

This thesis' goal was to craft a KM approach, that supports the inter-personal knowledge exchange by capitalizing on the currently possible benefits from contemporary approaches while avoiding their disadvantages. Therefore, following Becker's (2010) suggestion, in the descriptive evaluation this thesis' KM approach will be contrasted with the ones currently used but for which disadvantages could be found: *yellow page systems, expert recommender systems*, and *knowledge networks*. The evaluation will first be structured according to the *challenges* that were identified for the contemporary approaches in chapter 3 and that this thesis set out to remedy. Then the evaluation is complemented in accordance to the *advantages* that were identified for the contemporary approaches in the same chapter to ensure that this thesis' approach did not improve on the one side while losing on another side. In the end a *critical acclaim* of descriptive nature of this thesis' approach completes the descriptive evaluation. It works out problems that are not related to the other approaches but that emerge due to the way in which this thesis' KM approach is set up.

7.1.1 Comparing Disadvantages of Contemporary Approaches with this Thesis' Approach

Section 3.2.3 analyzed the disadvantages of *YPS* and found the following:

Id	Description
Problem 1	Initial setup is resource-intensive as modeling is not straight forward
Problem 2	Too high effort for the individual to keep data up-to-date; low motivation of individual
Problem 3	Frequent changes incur either high cost or information is not up-to-date
Problem 4	Non-integrated; lack of work context
Problem 5	Creates one-on-one interactions, that are not well manageable

Table 7-2 ***Summary of problems that Yellow Page Systems face***
Source: Own Illustration in section 3.2.3 of this thesis

In this thesis' approach, *competence and context models* are created automatically extracting the knowledge worker's trails that they leave in IS log files during their work. Therefore, there is *no initial cost for modeling* (Problem 1). However, there is a *larger technical initialization* cost for the approach itself as it needs to be integrated into an IT landscape, while YPS are often stand-alone systems and hence cheaper to implement from an IS point of view. Likewise, in this thesis' approach there is *no effort for keeping the user model up-to-date*. This is done automatically as the logs of IS usage are continuously fed into the model. Additionally, the low motivation of individuals to maintain their competency model is not a challenge for this thesis' approach. There is *no personal effort for the maintenance* and the fear of being contacted due to the status of being an expert is reduced in this thesis' approach: the *expert may simply ignore requests without social punishment* due to anonymity and, by using contextual overlap as additional matching criterion, the requests that do reach the expert have a *higher likelihood to be relevant* because relevancy of interactions is contingent on the work context (Problem 2). In the same line of argumentation, frequent changes pose no problem, as the *update of the model is done automatically* (Problem 3). Further, this thesis' approach explicitly operates on task contexts and hence *is not context-free*. The integration into operative IS is possible as section 7.2 will show, also the integration of the user interface into operational IS systems is one strong demand of this thesis' approach and is also feasible as shown by the case study evaluations later in this chapter (Problem 4). Finally, a disadvantage of YPS is their focus on one-on-one interactions that are hard to manage. This thesis' approach, while building on dyadic communication, has mechanisms to *increase the number of directly affected individuals* and a means for *storing interactions in a retrievable way* (Problem 5). In sum, all problems that YPS face are addressed and remedied by this thesis' approach.

Section 3.3.4 analyzed the disadvantages of *ERS* and found the following:

Id	Description
Problem 1	Relationship of keywords and person not necessarily authoritative
Problem 2	Definition of expertness not with respect to context and current need but global
Problem 3	Expert selection often not supported; no concept for supporting asynchronous interactions
Problem 4	Non-integrated which necessitates change of media; very seldom consideration of user context
Problem 5	Creates one-on-one interactions, that are not well manageable

Table 7-3 ***Summary of problems that Expert Recommender Systems face***
Source: Own Illustration in section 3.3.4 of this thesis

ERS sometimes struggle with reflecting expertise by analyzing textual documents. This thesis' approach does not use textual representations as evidence for knowledge but *performed tasks are used as evidence for knowledge*. As tasks reflect what a user is actually working on, they reflect expertise in a more reliable way (Problem 1). Another challenge that ERS face, is their typical approach to model expertise without taking contextual factors into account. This is problematic because *expertise is always relative to a work context* and the need for expertise *emerges in a contextual setting*. By capitalizing on the task context as its fundamental principle of work this thesis' approach does not suffer from this problem (Problem 2). Further, there are only few ERS that also support expert selection. This thesis' approach on the other hand has an expert notification mechanism integrated that addresses those experts that are more *likely to appreciate being contacted*, that is automated and, that can be configured to *operate anonymously* (Problem 3). Only few ERS take the current work context of an individual into account, and virtually no ERS are (in addition) directly integrated into operational systems – their usage creates a media break. This thesis' approach on the other hand, *is integrated into operational systems* and, building upon task overlap at its core, *takes the current context of an individual into account* (Problem 4). As argued above for YPS, this thesis' approach has measures to increase the positive influence of dyadic one-on-one interactions to a larger group of benefiters (Problem 5).

Section 3.4.7 looked at the disadvantages of different forms of knowledge networks. The first one were *(classical) CoPs*, where co-located individuals work closely together. For them the following disadvantages could be determined:

Id	Description
Problem 1	CoPs are difficult to manage; they are a managerial paradox
Problem 2	May be misused for political reasons
Problem 3	Knowing about the existence of a community as an outsider is difficult
Problem 4	Creation of advantageous but not yet existing communities is difficult
Problem 5	No support for when and how to split into subgroups
Problem 6	CoPs are very limited in size
Problem 7	(classical) CoPs can only form where people physically meet

Table 7-4 ***Summary of problems that Communities of Practice face***
Source: Own Illustration in section 3.4.7 of this thesis

CoPs are emergent structures and as such blossom on their own and diminish on their own. On the other hand, to make them useful for an organization, guidance is necessary. As an effect, there needs to be a balance between not managing the emergent structure and still managing it somewhat to reap the benefits for the organization. This thesis' approach is different to CoPs as it *supports emergent structures and fosters interactions*, but on a *more formal level* than typical for classical CoPs. Therefore this thesis' approach is less affected by the managerial paradox, however, mainly because it *capitalizes on a different social process* (Problem 1). A CoP may be misused by departments for example if they try to place individuals in the CoP to influence its development towards the department's goals. This is an inherent challenge, that this thesis' approach cannot address either (Problem 2). Knowing about the existence of CoPs so that one may become a member in it can be difficult depending on the policies of the CoP and its legitimization within the company. In this thesis' approach knowing about networks is *explicitly facilitated* and hence the approach is not affected by this disadvantage. Yet again, the support system itself must be distinguished from the social fabric that lies beyond it and that may exhibit such problems (Problem 3). The formation of new CoPs is a non-trivial task, as it might be unclear if there is need for it and who would be eligible to join. Typically, some thought leaders activate their loose social ties to form a new CoP but this process is often not structured. This thesis' approach suggests a viable method for the *formation of networks based on contextual overlap* that can alleviate this problem (Problem 4). Similarly, there is no decision support for when to create smaller sub-units due to diverging focus of splinter groups within the CoP. In smaller groups this might not be a problem but in larger ones it might become one. In this thesis' approach there is explicit support for this process using *task overlap as crystallizing means for the sub-networks* (Problem 5). Finally, CoPs are, in their classical definition, very much focused on co-located members that build up strong relationships through frequent, also accidental encounters. This limits their size, as co-location limits spatial distribution of members and time constraints limit the number of individuals that one can maintain strong relationships with. This thesis' approach does not build upon strong ties, but on weak ties (if electronic networks of practice are resembled) or at least less strong ties (if distributed communities of practice are resembled) and is hence, by design *less cosntrained in its possible size* (Problem 5).

Distributed Networks of Practice are the second kind of knowledge networks that were investigated. They can be described as CoPs whose members are typically not co-located, that meet infrequently and that maintain less strong relationships than in classical CoPs. For them the following disadvantages could be determined:

Id	**Description**
Problem 1	High effort for regular face-to-face meetings
Problem 2	Harder than in local CoPs to convey individual advantage with respect to effort
Problem 3	Less ability for lurkers to just learn than in less anonymous settings
Problem 4	Larger size creates problems of maintaining sufficient levels of know-who
Problem 5	Knowing about the existence of a community as an outsider is difficult
Problem 6	Creation of advantageous but not yet existing communities is difficult
Problem 7	No support for when and how to split into subgroups

Table 7-5 ***Summary of problems that distributed communities of practice face***
Source: Own Illustration in section 3.4.7 of this thesis

In Distributed Networks of Practice, members are not co-located but to still keep the members acquainted, typically infrequently face-to-face meetings are arranged. Due to the geographic distribution, this may inflict substantial costs. This thesis' approach is designed to either resemble knowledge networks (see later for comparison of disadvantages) or Distributed Networks of Practice. However, if it does resemble the later, it can only support the "digital part" of the interactions. A true Distributed Network of Practice still has to arrange for regular meetings and hence this thesis' approach cannot remedy this disadvantage (Problem 1). Along with the spatial distance comes *social distance* which makes it harder for members to see their gain in a lively community. While this thesis' approach tries to alleviate this problem by *routing interactions to individuals that also benefit from them* for their typical tasks, the problem of social distance may not be removed with this thesis approach (Problem 2). In Distributed Networks of Practice, members typically know each other, maybe only vaguely, but at least they can associate an individual by a name. This reduced anonymity also reduces the ability to lurk without receiving social pressure and hence lurking is less possible. While this may seem a good thing, lurkers also "contribute" by keeping themselves informed and hence by keeping the company's ability to act on a high level. In this thesis' approach lurking is possible, depending on the *configuration of anonymity* and which kind of social fabric it should support (Problem 3). When Distributed Networks of Practice grow, the number of individuals in them may exceed the limit of members that one can keep track of. In this case, it becomes hard to know who knows what. In this thesis' approach this is alleviated as *interactions are established based on the work context automatically*, which removes the need to have the otherwise necessary know-who (Problem 4). The remaining problems 5-7 are, in their essence, the same that were discussed for CoPs in the paragraph before and are not discussed again.

Electronic Networks of Practice were the final kind of knowledge network that were inspected. In Electronic Networks of Practice individuals are relatively anonymous towards each other, typically have diverse affiliations, and have never and likely will never meet in person. For them the following disadvantages could be determined:

Id	Description
Problem 1	Limited ability of self-reflection may decrease contribution likelihood
Problem 2	Minimum level of shared language and norms apply but are not (openly) negotiated
Problem 3	Abundance of messages limits willingness to find requests than can be answered
Problem 4	Abundance of messages limits capability to learn from other's interactions
Problem 5	Knowing about the existence of a network as an outsider is difficult
Problem 6	Creation of advantageous but not yet existing networks is difficult
Problem 7	No support for when and how to split into subgroups

Table 7-6 ***Summary of problems that Electronic Networks of Practice face***
Source: Own Illustration in section 3.4.7 of this thesis

Due to the anonymity in the network, it is challenging for individuals to assess their (relative) expertise and as a consequence they may be reluctant to help others, although they might be able. In this thesis' approach, individuals receive requests for knowledge when they are *likely to be willing to help* and *likely to be able to help* and they may, by receiving a request, *understand their value for others* better. This alleviates the first problem (Problem 1). Another challenge in Electronic Networks of Practice lies in the lack of socially constructed norms and at the same time implicit norms that establish over time or that are expected by some members. Newcomers have a harder time to adapt to these rules. By matching those individuals that performed common tasks, the individuals have a *shared language at least to the extent that is necessary for the performance of the tasks*. However, implicitly established rules in social groups will always spur and the challenges that arise from this fact cannot be addressed entirely by this thesis' approach (Problem 2). Another problem of Electronic Networks of Practice can be their own success, i.e. the size that that they may reach. When the network is too large, the information overflow may *decrease an individual's willingness to actively look for requests*, i.e. pull mechanisms might start to fail. In correspondence, the learning possibilities from interacting with a knowledge seeker and the *ability to learn by observing interactions decreases* likewise. This thesis' approach takes care to *limit the information overflow* to only those messages that are *likely to be relevant for the individual*, while still leaving the *chance to observe all interactions in a structured way* if desired and hence this thesis' approach alleviates this problem to a large extent (Problem 3 and 4). The remaining problems 5-7 are similar in nature to the other kinds of knowledge networks, however more pronounced in Electronic Networks of Practice due to their typical larger size. The performance of this thesis' approach with respect to these problems in the first paragraph of this section have already been discussed and will not be discussed again here.

In sum, this thesis' approach is able to *address most of the challenges* of contemporary approaches that it compares to. However, some problems, like formation of implicit norms are so deeply rooted in how individuals in groups interact, that this thesis' IT-supported approach may not change this. Nevertheless, *this thesis' approach is affected by less of the inspected problems than any of the contemporary approaches*. In the following, the analysis focuses on determining if this thesis' approach is also able to leverage as much benefit as the respective contemporary approaches can, as only circumventing problems does not necessarily yield a better solution.

7.1.2 Comparing Advantages of Contemporary Approaches with this Thesis' Approach

Analyzing the benefits that YPS generate, it turned out that well-maintained expert directories are, from a conceptual point of view, an easy tool that still can generate major timesaving for finding potentially suitable interaction partners. While this thesis' approach is believed to *find better matches easier*, this thesis' approach has to cope with *higher complexity*. An additional advantage of YPS is their potential integration to HCM IS and consequently the exchange of information with professional development activities to support tasks such as project staffing or professional development. While the analysis did not go into too much detail on this aspect, this thesis' approach could support this integration as well. Aggregating the information contained in the execution of tasks onto a level that is more abstract but used commonly in skill management, the approach can even give *more concrete evidence of expertise*. A drawback is that there needs to be an *adequate mapping* that aggregates the atomic tasks that a knowledge worker performs onto the necessary level of abstraction – which is a field of research on its own – and it induces the problem of constantly keeping the mapping up-to-date.

Turning to the advantages of ERS, one finds that they typically use interfaces that resemble those of search engines when prompting a knowledge seeker for his request for knowledge. This is an advantage as many individuals are used to search engine interfaces and have a natural expectancy of how the system operates. This thesis' approach on the contrary does not use a search interface. Instead, it directly asks the user to enter a request and the search for suitable interaction partners is performed automatically. While the benefit of lower effort may be communicated to the stakeholders, they still operate with a *different interface that might potentially be more repelling* than a typical search engine-like interface. An important aspect of ERS is their ability to automatically collect information that is used as evidence for expertise. Once installed and equipped with information about how the extraction of relevant expertise information should be done, the system operates without much maintenance. This is similar to this thesis' approach where a similar advantage, after initial configuration, is to be expected – both approaches are quite *flexible by design*. This is also true for changes in the *typical terminology* used in a company as those will be reflected in (newer) documents and ERS automatically adapt this way. In this thesis' approach, tasks are the unit of analysis and those *tend to stay constant over longer periods of time*. ERS are also quite flexible in another aspect. Depending on the concrete implementation, they may use various sources, like various kinds of documents or communication means to extract evidence for expertise. In this respect, this thesis' approach is more limited as it focuses on the execution pattern of tasks. The approach may utilize more detailed information contained in the logs of the underlying IS, such as performance information or monetary values if those are available, but the *principle unit of analysis remains task executions* in this thesis' approach. The final advantage elicited for ERS was their non-obtrusiveness. They feed their expert models using information that is created during the normal work of knowledge workers and hence there is no (perceived) extra burden. This is also the case in this thesis' approach, where tasks are performed in a natural way and the information contained in the task execution is capitalized on *without the need of a knowledge worker to interfere*.

For each of the different kinds of knowledge networks distinct advantages could be determined as well. Communities of Practice have become a quite popular measure especially in larger corporations in recent years. One benefit they offer is the *controlled redundancy* of knowledge they create. Experts in the CoP learn from each other which is why in an intact CoP the loss of one member does not jeopardize the CoP nor does it harm the company as much as it would when a single outstanding expert leaves. This thesis' approach can perform similarly but with a serious limitation. In CoPs much learning takes place in *informal settings*, by-chance encounters and other frequent interactions. This thesis' approach focuses on *explicit encounters* and the missing face-to-face interaction *removes much of the learning potential* that makes a CoP so successful. Hence, by the very nature of this thesis' approach it is not possible to maintain a redundant network of experts as it is in CoPs. Further, CoPs can be a source of information about which knowledge the individuals of the company possess and hence about the pool of knowledge of the company itself. This also applies to this thesis' approach, but with an indirection. The approach may indicate what tasks can be performed and how well. This way, it is *indirectly possible to deduce knowledge* – certainly with less accuracy though. Another advantage of CoPs is their ability to further the profession that it is about by establishing Best Practices and spreading new knowledge among its members. This thesis' approach focuses less on the creation of new knowledge but rather on the *spreading of existing knowledge*, with the exception of *learning effects due to the interactions of individuals in problem solving situations* as also suggested by TMS theory (see section 2.3.2). Hence, this thesis' approach is not capable of supporting knowledge development capabilities as strong as it is possible in CoPs.

Distributed Networks of Practice have the advantage of extending the number of individuals that can be part of the community to a much larger number than is possible in "classical" CoPs. This is also true for this thesis' approach that may support the "digital" part of a distributed Network of Practice.

Electronic Networks of Practice extend the number of possible interaction partners to a virtual limitless number of individuals. Being able to resemble Electronic Networks of Practice, this is true also for this thesis' approach. Therefore, the approach is likewise able to support *getting answers to complex problems quickly*, getting access to *knowledge that is locally unavailable* and to *learn from the mistakes of others* as well as to *enhance one's own self-efficiency without social pressure* in an anonymous form of social interaction. Due to the context-specific matching of interaction partners in addition to the "normal" forum-like interaction this thesis' approach should be even more rewarding in these aspects.

7.2 Architecture Analysis

A thoroughly derived and adequate solution is worth little if technical problems hinder the implementation. Therefore, the implementation of this thesis' approach into in a live landscape of many IS has been analyzed. This way, it is also possible to find out about possible *performance implications*, *maintenance efforts* and to get a feeling for the *cost of implementation*. The implementation is later used in the case studies as well. In this thesis' approach the implementation is treated as part of the construct *organizational context* as the available infrastructure determines the specifics. This is also true for the maintenance efforts that depend on the hardware and software within a company (see section 5.3.4).

7.2.1 The Existing Architecture

To better understand the implementation of this thesis' approach in a *real-life setting*, the architecture of the target landscape for the implementation is detailed in the following. To cater for a *sufficient level of complexity*, that is typical for landscapes in lager organizations, a large scale IS landscape was chosen.

This thesis' approach was instantiated in an *SAP system landscape* consisting of *roughly 100 SAP systems* that are mainly used for research and education purposes. Most SAP systems in this landscape are ERP systems that span four generations of this software with the majority being of the latest generation. In addition, the landscape contains also special purpose systems such as SAP Business Intelligence systems and systems tailored for specific branches (so called Industry Solutions). Each SAP system hosts a number of *clients*, which are self-contained data sets capable of hosting separate companies or, in the setting at hand, research and educational institutions. In turn each client may have up to *thousands of users* that perform tasks in the system such as creating sales orders, managing production, creating customer inquiries, implementing program add-ons or (re-)configuring the base system to name just a few examples.

7.2.2 The Extraction Procedure

As the SAP systems support explicit as well as implicit process models they are PAIS and are therefore suitable data sources for this thesis' approach (see section 5.6.1). Each SAP system automatically writes logs that contain detailed information about *which functionality* of the system has been used by *which user* for *which amount of time* as well as the *resources consumed* during the task execution. These logs are normally used for performance evaluations but also for tracking system usage by the SAP system management component. As they contain information about who performed which task at what time and consequently in which sequence, they are suitable data sources for an instantiation of this thesis' approach (see section 5.6.2). As the data is suitable, an extraction method has to be implemented to complete the implementation of the *Data Acquisition Layer* in this SAP landscape instantiation of the approach. There are multiple IS and in principle cross-system processes are possible – in real-life SAP landscapes such a setup is very typical. Therefore, the reference implementation should support this as well. A natural choice is to aggregate the information in one place. In this landscape, all SAP systems may exchange data through a shared file system. This fact is used to extract all data entities from their respective SAP system, accumulating them on a central location in the shared file system to then upload the overall data set into a *central SAP system*. As it is a goal to improve the performance of the retrieval, a data warehouse system was chosen for this purpose. Additionally, the data is cleansed and transformed during the process of extraction. This is senseful as in its original format it also contains much system performance information like database connection times and CPU cycle usage counts that are irrelevant for this thesis' purpose and are hence removed. The whole process is depicted in Figure 7-1. Not putting too much stress on operational systems was a requirement for the *Data Acquisition Layer*. This is why data is extracted in a 24 hours rhythm at night. However, in this case there need to be measures to capture the context of a user that is within this 24 hour window and hence possibly not transferred to the central system yet. This was solved by creating an application that is installed in every SAP system. It can read the local system's performance logs and *extract recent task contexts*. For many purposes this recent context may be sufficient, in the other cases the computation additionally relies on the data stored in the data warehouse. Furthermore, it turned out that the *Data Acquisition Layer* must make sure that the data is captured in a way that allows a later display of the model with terminology normal to the users. The entities that are extracted contain *transaction codes*, a SAP-specific terminology of the *tasks that have been performed*. While for many users these codes are common due to their frequent use in SAP, during the extraction process additional information in the SAP system was used to amend the data with natural language expressions of the tasks.

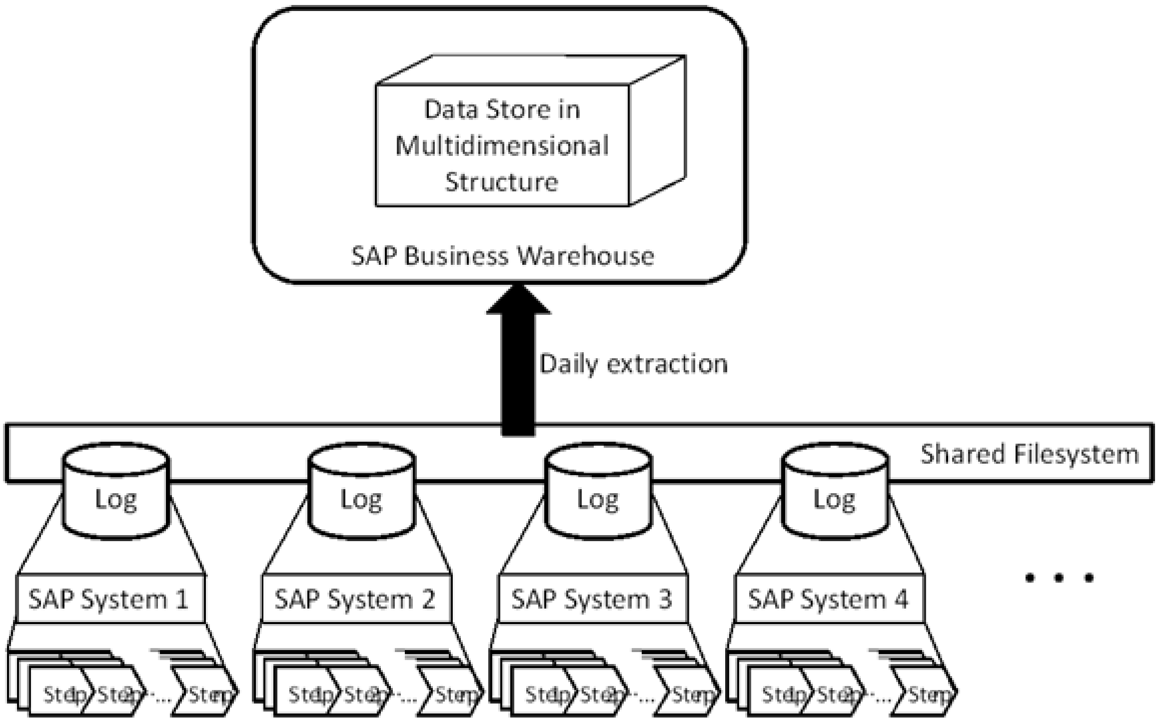


Figure 7-1 ***Data extraction in this thesis' target SAP landscape***
Source: Own Illustration

7.2.3 Results of Architecture Analysis

The architecture analysis' results indicate that the extraction, also in a *large-scale environment* with many systems and with many users is possible. Up to the day of writing, the collection mechanism has collected some 2 million tasks that were performed in the SAP landscape starting July 2009. The extraction process needed some tuning in the beginning and administrative costs were high. However, this was independent of the approach and could be resolved with tweaking some functionality in the SAP data warehouse solution. It now runs *unattended without administrative effort having average load times in the scale of minutes*. In sum, one can conclude that the demands of the *Data Acquisition Layer* are feasible in large-scale IT landscapes of PAIS.

7.3 Case study I – Small-Scale group

After evaluating the technical feasibility of this thesis' approach, in a next step a more important aspect takes center stage: real-life feasibility in terms of *user acceptance* and *possibilities for generating benefit*. It was chosen to do this evaluation in a two step procedure. First, a smaller group was presented a first *prototype* to learn about the general aspects of the approach, glitches in the implementation and general user acceptance. In a second, considerably larger field study the approach's possible benefits were determined.

7.3.1 Description of Setting

The goal in this setting was to gather feedback about this thesis' approach from users that have lived through the experience of *needing help while performing knowledge-intensive work* in an operative IS – the very situation that this thesis' approach was designed for. The group used for this purpose consisted of eighteen individuals (5 female, 13 male) with similar educational background. They were trained by the author in general aspects of data warehousing such as architecture of data warehouses, use of data warehouses, data modeling in data warehouses, typical administrative tasks in data warehouses, modeling and implementation of data extraction mechanisms (ETL – extract transform and load) as well as representation of data. The general aspects were subsequently illustrated in an SAP Business Warehouse (a data warehouse solution) before the group had time to try out exercises on their own. To reinforce their learning and understand their level of mastery of the system and the general design tasks, they were asked to individually do a small data warehouse project. It should include knowledge-intensive decisions that subsequently result in tasks performed in the prepared SAP system. A published case study (Hecht et al. 2008) was used for this step. Afterwards, the individuals were assigned to six different groups of three individuals each. Each group was tasked with conducting a data warehousing project that had an estimated complexity to occupy each team for six weeks. The project included designing a solution, blueprinting a solution and implementing the proposed solution in an SAP data warehouse system. Hence, the students had to perform knowledge-intensive tasks that resulted in using functionality of an SAP system. The individuals were shortly introduced to the KM approach as far as they were affected, orally in one session before the start of their projects and additionally by administering the Microsoft Word document of Appendix B-1 to each individual via e-mail.

7.3.2 Implementation

The implementation of this thesis' approach in the SAP data warehouse reused the *Data Acquisition Layer* implementation that was set up in the architecture analysis before (see section 7.2).

The *Mediation Layer* was implemented as follows. The user models consisted of task sequences that contained the following information: task (a *transaction* in SAP terminology), more specific part of task (a *function* in SAP terminology), executor (a *user* and a *terminal* in SAP terminology) and an execution timestamp (the start time of a transaction was used). Because the number of participants was rather small, a simple similarity measure for context overlap determination was used. If an individual creates a request for knowledge, it is sent to all individuals who have executed this task in the last 24 hours, i.e. someone is *similar if he has worked at the same task within this timeframe* otherwise not. The previously executed tasks were determined by reading directly from the operational IS without relying on the external data warehouse. This was sufficient as the individuals only operated in one system and the timeframe of 24 hours can be supplied by the operational system. The list of potential candidates was then ordered by number of times the task has been executed and the request for help was sent to the *three highest ranked individuals*. There is *no effort for the user*.

With respect to the *Representation Layer* not all functionality that was discussed for this thesis' approach was relevant for this setting. For example, since the approach was freshly implemented and the group was very small, *network size management* was not an issue and was not implement in the prototype used for this case study.

When a user wants to send a request for knowledge, he can do so by calling up a *specifically created transaction* in the system. Then, the knowledge seeker has to choose for which transaction he requests help. This is a small burden but since the SAP system allows concurrent connections from one user, this is necessary. Otherwise, it would not be obvious for which task help is needed, if more than one concurrent connection exists (see Figure 7-2, area 1). Upon the selection of the context the user could enter his request for knowledge (see Figure 7-2, area 2) and send it anonymously to an anonymous group. If a suitable user is currently logged on to the system he is notified by a so-called *SAP Express Message*, which is an *immediate pop-up*. If the user is not currently logged on to the system, the suitable interaction partner would receive a message in his SAP-integrated message inbox. In both cases, the *sender is anonymous* to the interaction partner. To increase the potential impact from mere dyadic relationships, the interactions are stored in a *list-formatted overview* (see Figure 7-2, area 3). There is a list overview for questions that the individual is engaged in, because he has asked this question and one for questions for which he was selected as a suitable interaction partner, and there is an overview for all questions that have been asked. All this lists are arranged in tabs that have similar graphical design and include, next to the message itself, the *context in which the request was created*, i.e. the task in which the need for knowledge emerged. There were no dedicated roles that users would assume, however, the author checked on the usage of the system and hence acted in role of a *coordinator*.

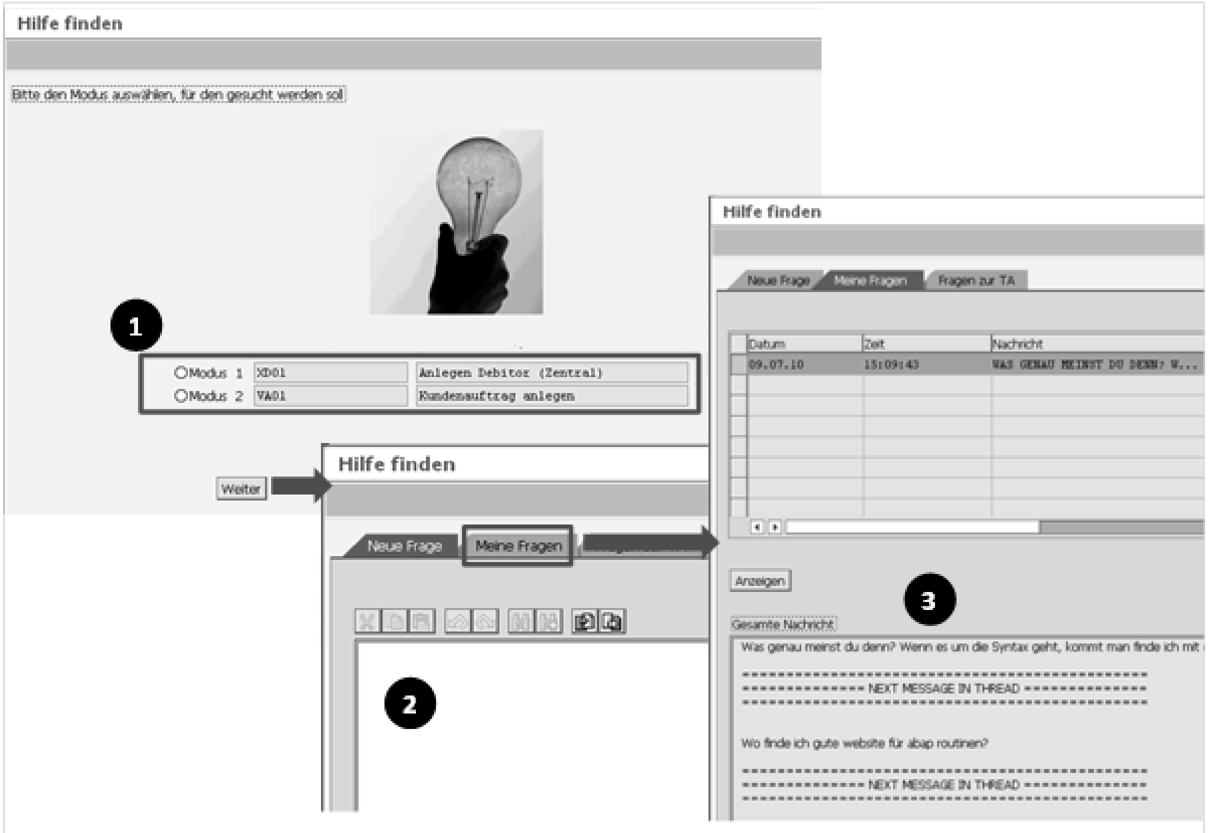


Figure 7-2 ***Overview of this thesis' first prototype's interface***
Source: Own Illustration

7.3.3 Result of Prototype Application

After the projects finished, the individuals were asked to *openly state their experience* with the application and return a digital version of their comments to the author. Since as much qualitative feedback as possible was to be collected, the individuals were asked the following open questions that were supposed to trigger elaborate answers:

- What do you think of the concept in general, what did you like about it, what did you not like about it?
- What do you think of the implementation of the concept, what did you like about it, what did you not like about it?
- Did you find any bugs in the software implementation?

The individuals were not incentified for giving feedback other than kindly asking for it. From the 18 individuals being part of the group it was possible to obtain five documents with feedback, that accounted for twelve of the individuals, as some of them returned group-wise feedback. This corresponds to a response rate of about 66,66 %. The qualitative results were extracted from the original documents, and for the purpose of including them into this thesis, were translated into English from their original language (German). They are collected in Appendix B-2.

Summarizing the feedback the following was mentioned once or multiple times:

- The individuals appreciated the *embeddedness* of the knowledge seeking application.
- The *reuse* of questions according to transaction was perceived useful.
- Taking into account the *work context* was perceived useful.
- *Anonymity* was perceived useful by a number of users but others would prefer the ability to choose, while still suspecting negative effects of anonymity.
- The individuals found *some implementation bugs.*
- The individuals missed functionality that they seem to be used to from typical forums, like *"closing"* a question and *responding to one's own question* or *deleting* it
- The individuals *felt limited* to only one system[10].
- The individuals felt that *existing solutions* (instant messaging) would not be replaced and that those might be faster means for problem resolution, especially since there is higher social presence[11].
- The *visual design* was sometimes fully appreciated and sometimes suggestions for improvements were made.
- *Processing speed* has been mentioned as too low.
- Individuals missed *further structuring components* like a title, or categories.
- Individuals missed an *incentive system.*
- Individuals felt that *deactivation by user choice* of the application should be possible.
- Some individuals wanted to *select whom to send* messages to.

7.3.4 Interpretation of Results

The qualitative feedback suggests that the group which was used for evaluation *liked the very idea* of this thesis' approach and especially *welcomed that it is embedded* in the system. Furthermore, *reuse of previous knowledge* was appreciated. Some implementation-specific remarks were made. For example, bugs were found, the user interaction was not well-accepted for some users, while for others it was, the limitation to one system was felt as a hindrance,

[10] The individuals had to use another SAP-specific program, that is technically very different. As the main system was the SAP data warehouse, an adapter for the external system was not implemented

[11] A comment like this is not suprising, as the individuals knew each other for an extended period of time, i.e. they had built up strong ties that they reinforced via media with high social pressence like instant messaging systems

processing speed was perceived too low and some additional functionality for structuring was asked for. In addition the individuals made some valuable remarks how to improve the implementation. For example, the provided functionality should *more closely resemble a forum* or the inclusion of an *incentive system* relying partly on *extrinsic incentives* was suggested.

From these results one can conclude that the approach itself is perceived to be beneficial but also some of its limitations became apparent. The implementation-specific imperfections identified in the first case study were used to improve the implementation of the approach which was then applied in a second evaluation cycle described in the next subchapter.

Apart from the implementation-specific feedback, it was also possible to learn about the viability of the approach on a more general level. Some of the suggestions stated that other viable sources like the SAP-specific help web page exist, others pointed out that they already have established means to transfer knowledge in their strong tie network using tools such as Skype, while still others emphasized the importance of a speedy answer to a knowledge request in urgent cases, that cannot be guaranteed with the application. Summarizing these comments, the implementation of this thesis' approach is essentially *seen as a communication channel* and hence the *social presence* and the *media richness* (see subchapter 5.2) of this communication channel are important. More specifically, the relative social presence and media richness *in comparison to already known and utilized communication channel* seems to influence the perceived added value of this thesis' approach.

7.4 Case study II – Larger-Scale group

Taking the feedback of the first evaluation cycle into account the next instantiation of the approach was refined as was its implementation. The new instantiation was subsequently applied in a second case study. For this second cycle of evaluation a different set of individuals in a somewhat different contextual setting was selected. The first difference in the setting related to the size. In the second case study a *much larger group* was addressed. This makes sense as the larger the group that may benefit, the higher the possibility for task overlap and hence for knowledge transfer between individuals. A second aspect that is quite different in the second setting is the *degree of acquaintance* within the set of participants which was much lower. This makes sense as it turned out in the first case study, that strong ties are preferred over the help that this thesis' approach may provide. With geographically dispersed individuals that *do not know each other*, the potential benefit of this thesis' system is much more pronounced. In addition, this setting is very similar to one found in a typical, large organizations.

7.4.1 Description of Setting

The target group of the evaluation again consisted of individuals that had to perform knowledge-intensive work and, as part of this work, they experienced the need to seek knowledge. The participants were tasked to model and configure a company in an ERP system, in this case an SAP ERP system. This task is challenging as there are numerous choices supported by the system with respect to the setup of a company. The participants belonged to one of two groups. The first consisted of less-experienced individuals that performed more general configuration tasks – the introductory level group. The second consisted of individuals who had mastered the general configuration tasks already, now being tasked to address advanced configuration challenges. Each individual worked independently on the same tasks of the respective group. Additionally, with only very few exceptions, the individuals *did not know each other in person*, as they were geographically dispersed across Germany. The groups consisted of 93 individuals altogether, 63 individuals in the introductory level group and 30 individuals in the advanced level group. The individuals were introduced to the KM approach as far as necessary for them and also the usage of the implemented application was described. The information was distributed to the individuals sending an e-mail to them that also contained the descriptive PDF document represented in Appendix C-1. This e-mail was sent after the individuals had started to work on their projects, almost exactly three month before the results of their efforts were due.

During the course of the case study, the usage of the approach's instantiation was observed and the author kept an eye on possible problems to keep possible disturbance due to software bugs and the like to a minimum. Also, listening to participant feedback an update was released during the course of the case study that contained minor bug fixes as well as slight improvements of the design that were asked for during the project phase. The participants were informed about the new features in another e-mail that again contained an amended PDF document that is outlined in Appendix C-2.

Additionally in this setting, there was another stakeholder. The group of individuals had the possibility to post requests for knowledge to 1st level support staff that was experienced in SAP customization and was acquainted to the specific tasks. The request for knowledge would be posted in a password-protected forum setup specifically for this purpose, where the 1st level support staff could then respond. Therefore, the 1st level support is another stakeholder in this setting. They naturally had an interest in reducing their workload and hence embraced the new solution that could move some of the effort back to the community itself. Quick response times to the knowledge seekers' inquiry was important to them as it defines their service level. The forum provided an established mechanism for this purpose and was therefore not replaced but rather amended by the implementation of this thesis' solution.

7.4.2 Adaptations to Implementation

Informed by the results of the first case study, this thesis' approach and its implementation were adapted. Although the user feedback was incorporated, the general approach as described in depth in chapter 5 *remained unchanged*. Only *the implementation-specific traits* of this thesis' approach were adapted. These implementation-specific traits have also been identified as critical factors to take the envisioned support towards providing operational benefit (see chapter 4 for the derivation of those factors). Specifically, the following categories of features were altered. They are also described in more detail in the following:

- Incentivation
- User interface
- Increased functionality for interaction
- Personal configuration
- Administrative features
- Improvement of backend functionality

7.4.2.1 Incentivation

One feedback of the first evaluation cycle suggested to use *extrinsic incentives* to spark motivation for participation. Considering the discussion of *motivation* being a critical success factor (see section 4.2.11) the possibilities for sparking motivation were carefully reevaluated. *Extrinsic rewards* tend to replace the influence of *intrinsic motivation* and are therefore often seen as counterproductive, however, in the initiation phase they can help increase contributions until high value content is available (Oldigs-Kerber 2007, p. 69). Therefore, still the *personal benefits* the participants may gain are emphasized by phrasing the potential benefits in the description of the implementation (see Appendix C-1 and Appendix C-2). Nevertheless, the participants were also informed that the most active ones will be rewarded with *non-monetary, material prizes*, which in the case study were books on the topic of the tasks.

Additionally, to further motivate the participants to experience the possible benefits of this thesis' approach during the phase of project execution a slight adaptation of the implementation was used. An additional *ranking* which is depicted in Figure 7-3 is introduced. It shows the *number of contributions* and *amount of positive feedback* of the individual in absolute values. In addition, it shows a comparison to the *group average* of those two key figures. In correspondence to *Social Comparison Theory* (Festinger 1954), the expectation is that the participants feel obliged to be *above average* and hence *contribution motivation* is increased.

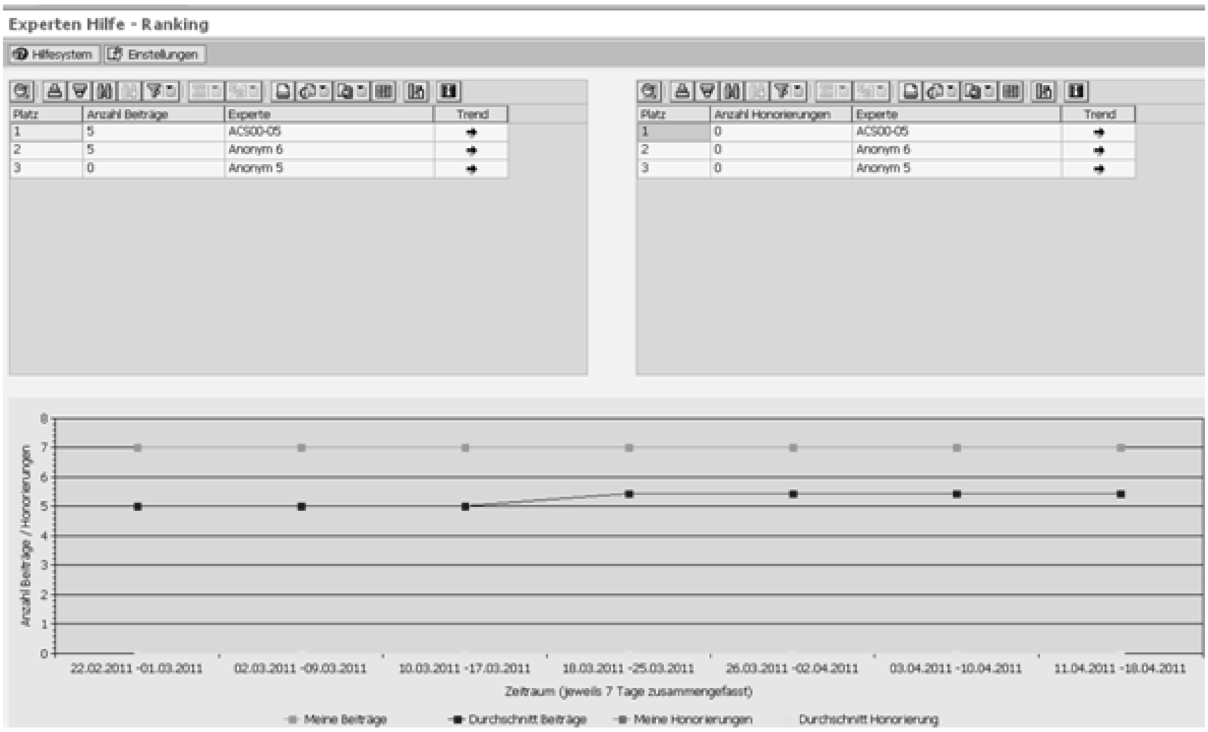


Figure 7-3 ***Ranking view in the updated version of the second prototype***
Source: Own Illustration

7.4.2.2 User Interface

Some of the feedback from the first evaluation cycle, suggested adaptations to the user interface, while others felt it to be well-organized. A design that is perceived as adequate is a crucial factor for any IT system and for KM application in particular (see section 4.3.9). Therefore, the general structure of having a section for questions that an individual has posted, another one for questions and a third one for the collection of questions was not changed. However, additional information was shown as well, such as the *number of answers*, the *status of the interaction* (whether it has a solution or not), and through which *interaction channel* it reached the individual (whether the individual was determined as an expert by the system or the user engaged into the interaction on his own). Also, the display of one interaction thread was improved to show the initial request for knowledge and the subsequent answers in an orderly way (see Figure 7-4).

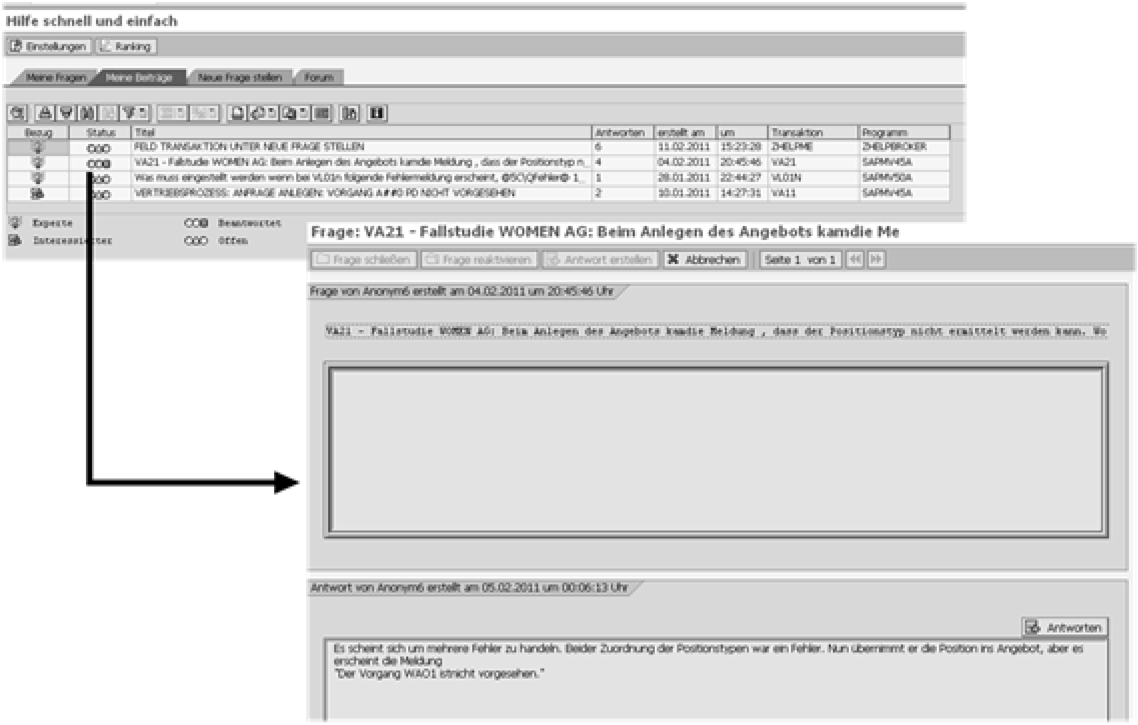


Figure 7-4 ***Display of list of interactions and one thread of interaction in the second prototype***
Source: Own Illustration

7.4.2.3 Increased Functionality for Interaction

Reviewing the feedback of the first prototype, another area for possible improvements is an extension of interaction possibilities with the system to deliver a richer user experience. Along this line, an increased ability to personally structure and evaluate knowledge pieces has been found to be a critical success factor for KMS (see section 4.3.6). To address this aspect, functionality to *judge the value of an answer* was included. The individual posting a request for knowledge can indicate for each response, if it is a *(partial) solution* to the request and can additionally indicate *positive feedback* and hence reward the contribution of others (see Figure 7-5). What's more, the initial requester can also indicate a thread to be closed and hence indicate that no more interaction is desired.

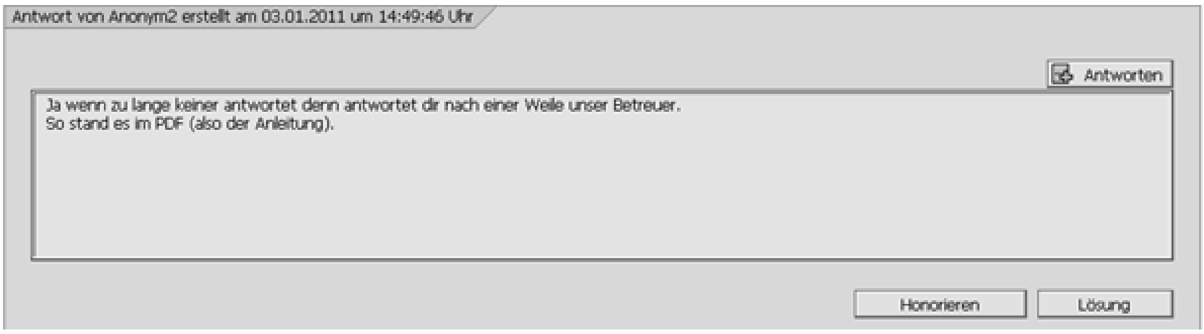


Figure 7-5 ***Interaction possibilities on a response in a thread in the second prototype***
Source: Own Illustration

Furthermore, to increase *knowledge reuse*, a supportive facility was implemented that presents *similar interactions* to a knowledge requester based on the current task, for which the requester needs support (see Figure 7-6). When clicking on one of the similar interactions the requester may check if it fits into his context and may either be informed sufficiently already or may restart the interaction in this request by posting a response himself. If the previous interactions are not suitable at all, the requester may still pose an own new question. This allows more *easily navigating existing knowledge* which is helpful for KMS (see section 4.3.7).

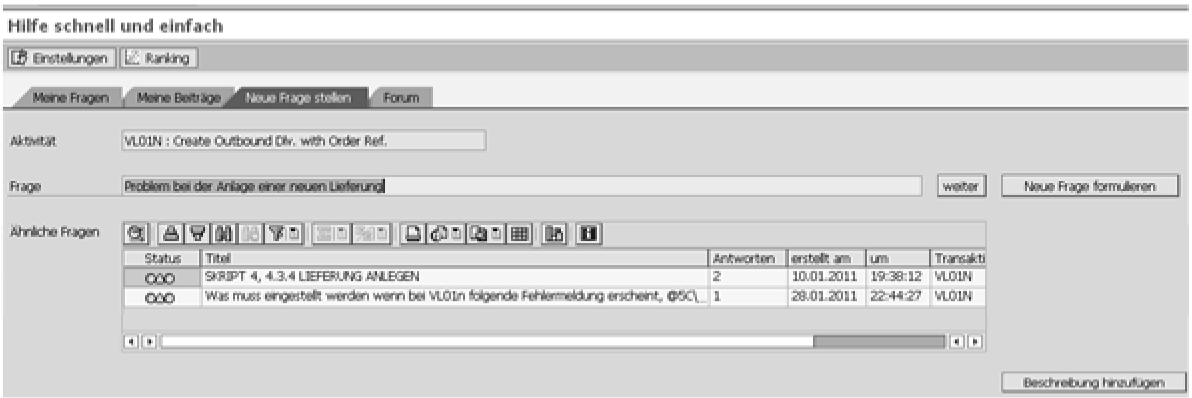


Figure 7-6 ***Support for reusing interactions in the second prototype***
Source: Own Illustration

Finally, another interaction possibility was included. When a knowledge worker is in need of specific knowledge contingent on his current task, he can still use the possibility to post a request for it. To make past experiences of others better accessible and to allow for *serendipity*, i.e. "accidently" stumbling across possibly relevant interactions (Roberts 1989), also a collective view was integrated that not only shows the interactions for one task but for all, organized by task in a tree structure (see Figure 7-7). This view is structured according to the terminology that is natural to the participants, i.e. their tasks and hence provides a *suitable knowledge structure* which is beneficial for the success of a KM approach (see section 4.2.8).

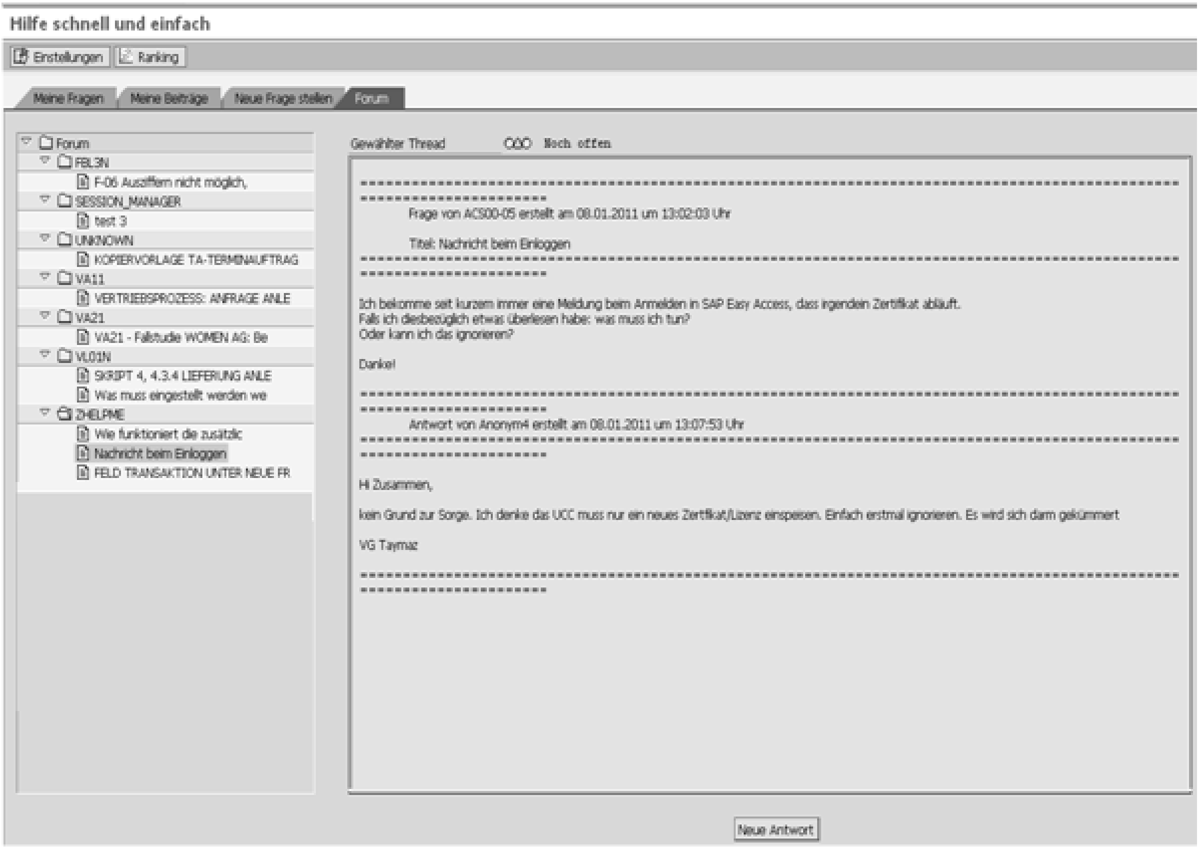


Figure 7-7 ***Structuration of previous interactions using concepts familiar to the users in the second prototype***
Source: Own Illustration

7.4.2.4 Personal Configuration

In the general KM approach it was argued that different *degrees of anonymity* may be useful and complete anonymity was established in the first prototype. The feedback of the first evaluation cycle suggested that some individuals of the target group might appreciate the ability to remain anonymous, while others might be more inclined to have their name publically visible. Properly treating anonymity is an important factor for the successful adoption and use of KMS (see section 5.1.2). One important motivation factor for not being anonymous is the possibility to be positively associated with one's contribution behavior and hence to increase the *(generalized) social capital* within the group of participants or more specifically to increase *reputation*. To include the possibility of being anonymous and still being *reciprocated with reputation* the participants were given the deliberate choice of configuring their user to be anonymous or to show their identification, i.e. user name. Their user names, which could be traced back to their real life name by a third party, is shown, whenever the individual explicitly sets this option, otherwise an automatically generated ID is shown whenever the user poses a request or enters a response to a knowledge request. Additionally, the individual may choose to take part in the ranking view (see Figure 7-3), depending on the previous configuration item, either with the user name or with an anonymous ID.

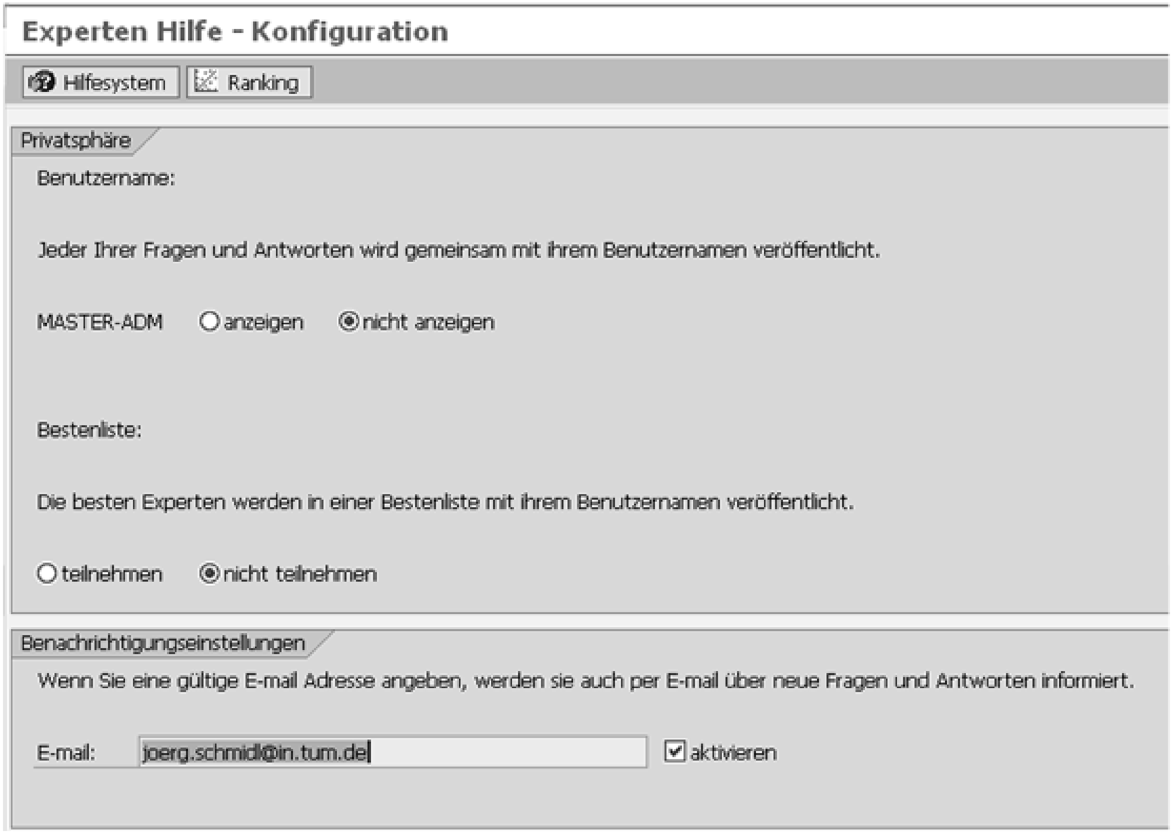


Figure 7-8 ***Configuration options in the second prototype***
Source: Own Illustration

Another configuration option that was included in the second prototype relates to the *preferred communication medium* for new messages. Messages are sent to an individual, if he is determined as *expert for a new request* or, whenever a new interaction is posted within a knowledge request that the *individual has previously participated in*. The individual may choose to enter his e-mail address in the configuration screen. In this case he is automatically also informed via e-mail. If this option is not used, only SAP-internal messages are distributed. The two communication channels induce different levels of *social presence* for the participants which is an important trait for systems that support interaction among individuals (see chapter 5.2). All configuration possibilities can be seen in Figure 7-8.

7.4.2.5 Improvement of Backend Functionality

Some of the feedback of the users related to the processing speed of the KM approach's instantiation or more specifically, they related to the perceived speed of the interaction with the application. Therefore the focus for *improving the user experience* was put on finding ways to improve the part of the processing that is visible to the individual. As a result the processing of the matching was redesigned. Now, the individual may post the request and immediately continue to work, while in the background the somewhat time-consuming determination of a suitable candidate is performed asynchronously without blocking the user interaction. This is a relevant adaptation as it increases the users' *perceived ease of use*, which according to the *Technology Acceptance Model* (Davis 1989) increases the *intention to use* the KMS. Analogously, reduced wait time decreases the users' *effort expectancy*, which according to the *Unified Theory of Acceptance and Use of Technology* (Venkatesh et al. 2003) has a positive influence on *usage and adoption* of the KMS.

7.4.2.6 Administrative Features

As noted before, in this evaluation cycle another stakeholder needs to be included into the consideration: the 1st level support who helps the participants with problem resolution and focused guidance on task-specific questions. Their goal is high user satisfaction that is achieved among others through low response times to the participants' requests. From their point of view, the instantiation of this thesis approach is also another communication channel for support requests. Therefore, to increase acceptance and use for this stakeholder group some additional functionality was included. Whenever a participant posted a request for knowledge, previously defined *knowledge stewards* also receive this request independent of their work context and their level of experience. They incorporate a *supportive role* as discussed in section 4.2.6. To further support the individuals assuming this role, a mechanism was additionally included that informs in regular intervals about knowledge requests that have not been addressed by an expert at all. This is an important measure, as especially when a new KMS is implemented, having quality content and quick response to user needs is of utmost importance to step into a self-reinforcing cycle of user satisfaction with the system and subsequent own contribution to it (see section 4.2.11). Finally, to better understand the users' behavior a final addition to the implementation was made. The users' click streams within the implementation were recorded to understand their search behavior.

7.4.3 Results of Prototype Application

After the final deadline of the participant's projects, the usage of the implementation of this thesis' approach was evaluated and the benefits and drawbacks that the participants experienced were determined. This was achieved using two different means for gathering information. First, the usage of the application was inspected using *log information* that was captured during the use of the implementation. This way a better understanding of real usage behavior as reflected in the system usage is possible. Second, an *online survey* was distributed to the participants. It contained questions relating to *assumptions and design choices* motivated by theoretical models that were used in the design of this thesis' approach and its implementation in the SAP system. Additionally, the survey contained items that asked for the participant's feelings towards *usefulness* of the implementation and its *handling*. The results back the usefulness of this thesis' approach as detailed in the following.

7.4.3.1 Application Usage Behavior

In this thesis' implementation a mechanism was added that writes a log entry every time the GUI sends an event to the SAP server. This allows to capture much of the interaction with the application. However, some aspects such as changing between tabs, are processed on the frontend only and are not captured. In sum 453 atomic interactions[12] were captured in the time frame between 21.12.2010 and 18.03.2010 that resulted in twelve questions asked and nineteen answers given.

Having a look at the aggregated results of the log analysis, showed that some functionality was used more often within the group than other. Figure 7-9 shows that the *forum and its functionality were used most often*. This is consistent with the assumption that many participants are lurkers (Wenger et al. 2002) and supports the contribution of benefit of this thesis' approach, as one design requirement was to integrate interested lurkers into the knowledge transfer. The second most used functionality was switching between fundamental parts of the application, i.e. between the *ranking display*, the *configuration options*, and the actual application. This indicates that ranking and configuration were important functionalities for the participants. The third in frequency of usage was functionality to *view an individual request* for knowledge exchange, hence, checking on the progress of a knowledge request. This is closely followed by functionality to enter new questions.

[12] This represents the filtered number, i.e. interactions that were caused by the 1st level support or by the author have been removed to only reflect the participants interactions

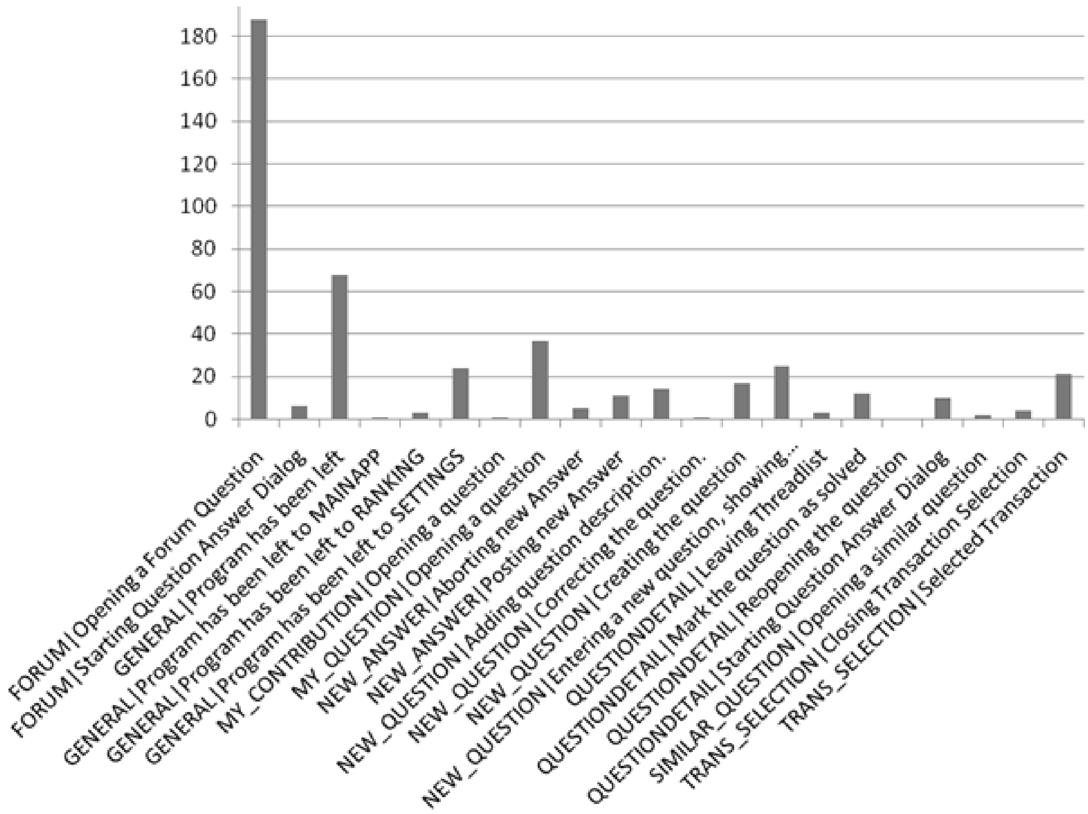


Figure 7-9 ***Distribution of number of times that different functionality in the implementation was used***
Source: Own Illustration

This implementation of this thesis' approach was used constantly during the case study's lifetime. However, there were obvious peak times for example around 10.01.2011 and 11.02.2011 as can be seen in Figure 7-10. At these times, not only the general activity as reflected by the number of log entries was high, but also the number of questions and answers was higher at these times. This suggests that the more participants are active at the same time, the more they interact or put differently, *if there is lower temporal distance, interactions are more frequent*. This is in line with expectations.

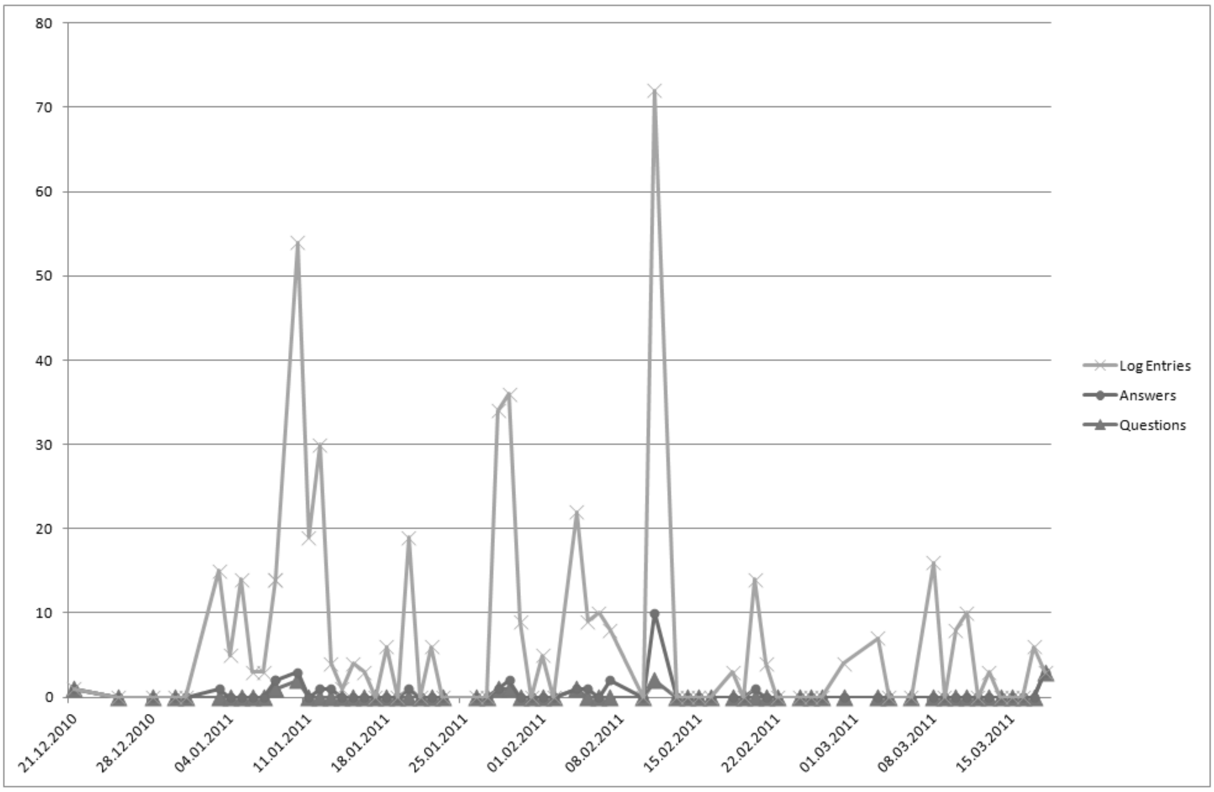


Figure 7-10 ***Distribution of number of log entries, answers and questions over time***
Source: Own Illustration

To analyze the interaction in further detail, the software package R[13] was used to aggregate the atomic interactions into interaction sequences. The procedure was as follows. Every atomic interaction was added to an interaction sequence as long as the time between the last added atomic interaction and the new interaction was less than 15 minutes. If the time between the last atomic interaction and the currently inspected one is more than 15 minutes the currently inspected one is considered the start of a new interaction sequence. In so doing 108 interaction sequences could be identified. The results show a distribution between very active participants (the first four in Figure 7-11), somewhat active participants (the next nine in Figure 7-11) and one-time users. On average the participants spent about 20 minutes in total using the application with a maximum usage of "User 4" who alone amounted to more than 2 hours. On average the participants spent a little more than 2 minutes on each interaction.

[13] R is an open-source data analysis software frequently used in academia. More information and download possibilty can be found at http://www.r-project.org/

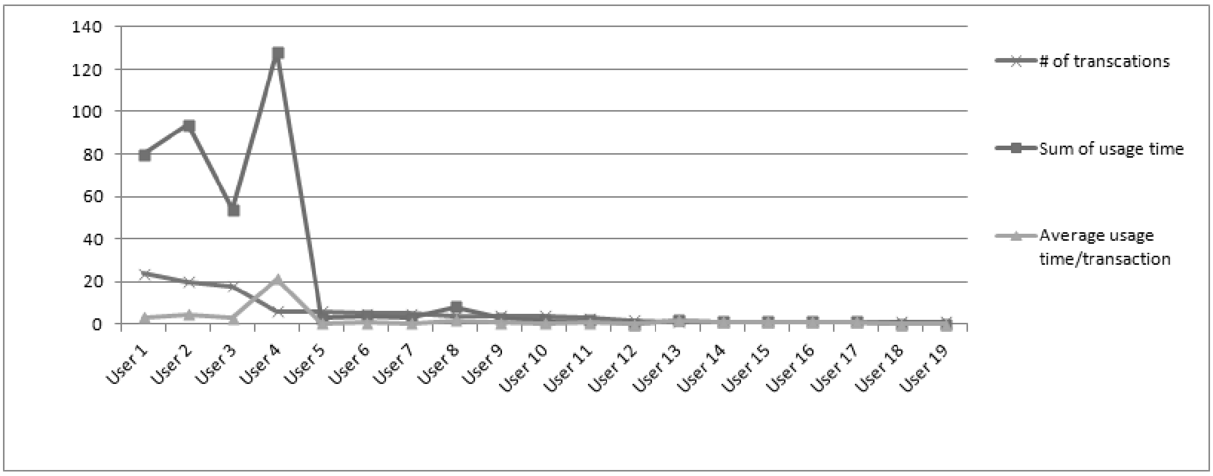


Figure 7-11 ***Distribution of number of transactions used, usage time and average usage time across all participants***
Source: Own Illustration

7.4.3.2 Survey Feedback

To validate the assumptions leading to the design of this thesis' approach and the implementation in this case study and to understand the benefit that the approach may actually yield, an online survey was distributed to the participants of the case study. It was constructed focusing on some aspects of this thesis' approach. First the survey checked if the participant was informed about this thesis' approach and if he understood the (user-visible) workings of the approach. Then, some items aimed at determining the participants understanding and feeling towards the term *expert* and the *message matching* that builds upon it. In a third part of the survey the *influence of anonymity* for the individual participant was elicited. The fourth part of the survey focused on the *interactions with the other participants* and also on *mechanisms that motivated* the participant to contribute. The next part of the survey focused on evaluating the *user interface* and its *integration into the operational system*. The final part of the survey asked general questions about the *usage of the application* and the participants' typical *information seeking behavior*. The whole set of questions as they were distributed to the participants can be found in Appendix C-3.

7.4.3.2.1 Survey Administration

Before sending the survey to the target group, it was *pre-tested* by asking a group of SAP experts that had not used the application before, to improve the clearness of the questions. As a result, some adaptations were made to make the wording more concise. Then, an invitation to participate in the survey was sent to all participants via e-mail. The *participation was incentified* by the chance to win one of six prizes namely SAP books that were relevant to the subject that the participants worked on during the case study. The survey was open between 06.04.2011 – one week after the participants' final project deadline – and 29.04.2011. About in the middle of this timeframe, on 17.04.2011 a friendly reminder was sent to further increase participation in the survey. Out of the 93 individuals that had the possibility to use the application, *21 returned completely filled surveys*, which corresponds to a return rate of 24,73 %. The goal was to gather feedback from those participants that did use the application and those that did not likewise. Therefore, the survey was created in a way that it branches in some parts to alternate parts. Depending on the survey participants' answer on whether or not they knew about and used the application, one version for participants that had no exposure to the application was shown or an alternate version for those that had exposure to it (see also Appendix C-3).

The results of the survey are detailed in Appendix C-4. Out of the 21 responders, three had no exposure to the application and hence filled out the alternate set of questions, the other 18 used the application, looked at it and/or read the documentation explaining the concept.

7.4.3.2.2 Items Relating to the Concept

The *application's concept was well understood*, only 11 % of the respondents indicated a neutral position to this question, the others gave positive indications of understanding. Additionally, the concept was *perceived to be senseful* with 94 % of the respondents giving positive or very positive indication. Also the structure that was employed – structuration according to task which in SAP is a transaction – was seen as helpful with 89 % positive or very positive indications to a corresponding question. Nevertheless, a third of the respondents were undecided whether finding previous contributions was easy and 6 % indicated that it was not. Also, in terms of the application's functionality, it seems that the *determination of expertness* was perceived to be only moderately accurate. Almost 80 % were undecided whether or not the determination was good, 17 % indicated that it was good and 8 % that is was not which suggests only moderate suitability. This is in line with the respondents' *expectation to be seen as expert* that returned similar responses. This result, however, matches expectations. The *contextual overlap* was overemphasized in contrast to *indication of expertness* as the number of participants and the typical interaction behavior as reported by the 1st level support staff suggested that there will be only *limited temporal overlap*. Hence the limit that indicates when someone is an expert was decreased to cater for this circumstance which ultimately increases the chance of indicating someone as expert when in reality he is not or at least does not feel this way. This interpretation of the responses fits the results of asking the participants for the *match of questions to their work context*. Here, only 17 % indicated that the match was not entirely suitable while 22 % indicated it was and 61 % were undecided. The contextual match also seemed to be important to the respondents as 11 % indicated that it strongly is, 33 % indicated that it is while only 6 % indicated that it was not and the remaining 50 % were undecided.

Further, the results of the survey suggest that for some respondents, being able *to influence the determination of receivers* is important. 23 % felt the need for this possibility, out of which 5 % did so strongly. On the other hand, for another 17 % this was not an issue and for another 5 % no issue at all, while the majority of 56 % was undecided if this was relevant or not.

Further, the approach's mechanism to *include other individuals* that might be interested in a discussion is in line with participants expectations. 22 % stated that they very much would like to be integrated in discussions relevant to them and another 56 % stated that they would like this.

7.4.3.2.3 Items Relating to the Interface

During the design of this thesis' approach the argumentation stressed that integration of the approach in the operational IS was important. This seems to match the participants' view. For 70 % of them the *ability to ask questions directly in the operational IS* was very relevant or relevant. Moreover, almost 50 % of the participants stated that they would *post more questions* if they would not have to change media. The implemented interface seems to be well accepted. 56 % found the interface easy or very easy to us and only 5 % indicated that they did not find it easy to use. Similarly, 50 % got along very well or well in the interface with some 17 % stating to have had some trouble. Nevertheless, the general navigation concept was perceived to be concise by 47 % and very concise by 12 % with an additional 41 % being indifferent to it being concise or not.

7.4.3.2.4 Items Relating to Decreased Effort

The design of the approach was driven by the idea that a similar work context would decrease the effort for request and answer formulation. The results of the survey seem to support this assumption. 33 % indicated that *formulation of requests is easier* and 11 % even indicated that this is very much the case, when they know that the receiver of the question is in a similar work context. Only 6 % did not think that this is case with an additional 6 % strongly not believing that it reduces effort and the remaining 44 % being undecided. The results for the *expectancy towards decreased effort for formulation responses* was similar but even more in favor of the assumption that effort can be decreased by this thesis' approach. Only 5 % thought that context overlap does not facilitate formulation of responses, while 17 % strongly thought it did and another 28 % thought it did with the remaining 50 % being undecided. The *information overload* that this thesis' approach tried to limit appeared to be no problem for the participants. 28 % indicated that they strongly felt it to be no problem and another 33 % believed it to be no problem. This is in line with the results of another item that asked if the users rather browse for relevant messages or if they want to have them send to them automatically. Here 67 % stated that they strongly favor or *favor to automatically receive messages* that the implementation of this thesis' approach deemed suitable for them. The general idea of automatically receiving messages was perceived useful by 61 % and perceived very useful by 11 %. Obviously the participants *did not fear to receive unnecessary messages* when not in control of the distribution mechanism. Also 53 % of them liked the mechanism that *sorts the messages automatically* so that they do not have to categorize them on their own.

7.4.3.2.5 Items Relating to the Participants' Motivation

Different factors were suspected to influence the participants' motivation to contribute. In contrast to the expectation, *being seen as expert was valued quite differently* among the participants when contrasted to the willingness to contribute in other communication media without being considered expert. 22 % of the respondents indicated that being determined as expert does not or not at all extend their willingness to contribute, 50 % were undecided as to its influence and 22 % indicated it does increase their willingness and another 6 % indicating that it strongly influences their willingness. This is in line with the respondents' answer to the question whether *others seeing them expose their expertise is relevant* to them, where only 5 % indicated that this is very much the case and 17 % indicated that it is the case. On the other hand the majority of 57 % were undecided and another 5 % did not see an influence and 6 % indicated that there is no influence at all. Yet, the *influence of contextual overlap was quite pronounced*. 61 % of the respondents were undecided whether or not they would respond more if requests sent to them match their context, while a third indicated they would respond more and 6 % even indicated that they strongly think they would contribute more when the request's context matches their work context. Also *receiving feedback* was important to the participants with 61 % indicating that it is important and another 11 % strongly indicating so. Only 6 % stated that receiving feedback is not important and 22 % remained undecided. Being considered an expert was stated to increase the responsiveness for 61 % out of which 17 % indicated that this strongly applies. With the remaining 39 % being undecided, the results suggest that *the feeling of being seen as expert is a strong motivator to contribute*. Seeing that *in comparison to others a participant contributed more motivates* 44 % to contribute further, with 6 % indicating that this is not the case and another 6 % strongly indicating that this is not the case. Surprisingly, however, a *ranking* that visualizes this difference was stated to not motivate further contribution for 33 % of the participants with 11 % of them even stating that they strongly think it does not motivate them. It was also interesting to see that the *extrinsic motivation* implemented by offering SAP books to the most active participants only motivated 28 % of the participants while 16 % stated that this did not motivate them and another 28 % stated that it did not at all motivate them. Furthermore, the results seem to suggest that the participants are *motivated by expectation of generalized reciprocity* as 17 % stated that they definitely expect to be helped also later with an additional 56 % stating that this applies for them also.

Personal benefits also were strong motivators for the participants. 50 % indicated that they strongly agree to receiving questions being a good way to learn SAP with another 39 % that indicate that this statement applies for them. Similarly, 39 % strongly felt that *interactions with others allowed them to learn* about SAP better with another 55 % that felt so as well. Surprisingly, the participants were also very altruistically motivated as 78 % stated that it is *important to them that others can learn from them* with another 17 % that strongly agreed.

7.4.3.2.6 Items Relating to Anonymity Effects

The effects that anonymity would have on the participants were also a relevant part of the survey. Apparently for a larger part of the participants *the concept of anonymity was important*. Asking anonymously was important for 33 % and another 9 % indicated that it was very important, while on the other hand 24 % indicated that it was not and 5 % that it was not at all important with 29 % being indifferent. Similar results were returned when asked for anonymous responses. This seems to suggest that for more than a third of the participants the feeling of anonymity is important. A similar distribution can be found when looking at the responses to the question whether the participants would ask more when anonymous and whether they would ask questions that they otherwise would not. Apparently, for roughly 40 % of the participants the *anonymity had a positive influence on posting requests*, which suggests that anonymously asking questions was relevant in this case study. When looking at giving answers anonymously however, only 5 % indicated that anonymity increases their willingness to respond more often while 14 % indicated that this certainly is not the case and another 24 % that it does not apply for them. At the same time 29 % indicated that they would post responses that, without being anonymous, they would not with another 43 % undecided if they would respond more or less and 9 % and 10 % indicating that they would not or certainly not respectively. This suggests, that *anonymity has only limited impact on increasing the number of responses* but for approximately a third of the participants it allows them to contribute without fear of disadvantage and hence *increases overall contribution* within the group.

7.4.3.2.7 Items Relating to the Participants' Background

Most participants (90 %) were older than 26 years of age, which, given that all of them have a university education, suggests that they are well educated. The participants also appeared to be rather *proficient in the use of SAP*. 95 % of them have used SAP before and on average they have attended 2,7 other SAP courses before. Also they indicated that they had a good feeling for whether or not they are experts in certain parts of SAP, with 44 % being undecided, yet 39 % indicating that they may judge their expertise well and 17 % indicating they could do so very well. This also reflects in the participants' perceived complexity to formulate suitable requests. 61 % stated that it was *easy or very easy for them to formulate questions* and only 6 % stated that it was not easy for them. As was expected, they *did not know each other too well* with 39 % strongly indicating that they did not know many other participants and another 33 % indicating the same. Only 6 % indicated that they knew many other participants with 22 % indicating acquaintance with others to some degree. The participants *were also used to interact with others* as 53 % of them have posted at least one question, 19 % of them even more than three. However, when comparing these numbers with the logs of the application, at least some of the participants apparently interpreted the question to cover all communication media available to them and not only the instantiation of this thesis' approach. The same holds for the results to the question how many responses the participants had received that showed similar values.

The participants also had a *high level of trust in the group*. When asking others 44 % strongly agreed that they did not expect any disadvantage and another 39 % agreed as well. When helping others half of the participants strongly agreed that they fear no disadvantage with another 22 % sharing this feeling.

Before the introduction of the implementation of this thesis' approach into the group of participants, other means of interaction had been provided for the participants to interact and those were also actively used by the participants. This is why the survey also tried to find out about the participants normal choice of communication media. 33 % of them very often use internet forums, 19 % do so often and another 24 % do so from time to time, while only 24 % stated to use internet forums seldom. This suggest that *the concept of forums is natural to the participants*. However, when having a question 43 % of the *participants rely on e-mail to find help* and another 33 % use the course forum, the remainder using either public forums (14 %) or personal meetings (5 %) and other means such as Google internet search (5 %). This answer pattern was expected as e-mails to the 1st level support and use of the course forum are the predominant means of support that was advertised to the participants. What is interesting however, is the participants' perception from where they received the most valuable responses. Although e-mail and the course forum were established communication channels almost half of the participants indicated that they also *received valuable answers through the application* that implements this thesis' approach. Also the *time until getting an answer was comparable to that of e-mail interaction*, however the forum was named more than twice as often for being a quick way to receive answers. Most likely this is due to the fact that this medium was more acquainted to the participants and, more substantially, that the 1st level support monitored the course forum to give speedy feedback while this was not the case for the application implementing this thesis' approach. It lived without organized support staff.

7.4.3.2.8 Overall Impression of the Participants

While checking on the benefit of this thesis' approach with diverse questions as described in the paragraphs before, also more direct response from the participants were considered relevant. Therefore, the participants were asked how they perceived the Expert Help. A reassuring *33 % of the participants perceived the application to be very good, another 45 % as good* and the remaining 22 % were undecided. Also the free text comments that the participants could enter if they liked, indicate that the application had the potential to support the participants. Some simply stated "great tool"[14], others indicated that they would use it again[15] while still others liked the concept although they did not use it or not extensively at least[16]. Nevertheless, some participants also made suggestions for improvement. Some suggested to improve search capabilities[17] or better documentation[18].

7.4.4 Interpretation of Results

Based on log analysis of the usage behavior, the application was *used throughout the case study time* with certain peak times of usage. The log also indicates the activity level of participants varies with some very active participants, others somewhat active and also a larger amount of non-active participants, which is in line with prior research about the composition of knowledge networks (Wenger et al. 2002, p. 56). The forum functionality was used most frequently, indicating that *learning from the interaction of others played an important role*. Furthermore, the survey results indicate the *usefulness of this thesis' approach*. Free text statements of the participants underline their appreciation of the concept, as do results of questions directly aimed at eliciting the satisfaction with the application. Most of the assumptions that informed the design of this thesis' approach were backed by the answers of the survey participants. Most importantly, the results support the fundamental concept of including *contextual overlap* in the determination mechanism of suitable interaction partners. However, determination of expertness was felt to be only moderately accurate. Given that the participants had only moderate temporal overlap, i.e. did often not work at the same times and their relative equality in expertness, this comes however with no surprise.

Relating to the SCT model (see section 5.2.1) that was used as one of the pillars for the theory-driven design of this thesis' approach, most of the constructs appeared to play a role

[14] Original comment in German: „klasse tool"
[15] Original comment in German: „Werde ich nochmal nutzen"
[16] Original comment in German: "Ich fand das Tool klasse. Da ich keine Probleme bei dem Kurs hatte nutzte ich es allerdings (leider) wenig"
[17] Original comment in German: „Manche Punkte finde ich ausbaufähig, wie Suchhilfe/-möglichkeiten. Hat mich etwas irritiert. Aber ansonsten eine klasse Idee!"
[18] Original comment in German: „Die Expertenhilfe war am Anfang etwas schwer zu bedienen. Vielleicht sollte die Anleitung mit einer "Schnellstart-Anleitung" erweitert werden: Wie stelle ich eine Frage? / Wie antworte ich?"

for the participants. *Reciprocity* was important for the participants and when asked as expert they stated to contribute more, i.e. they *felt committed*. They also judged their *ability to rate their expertise* as relatively high and had *comparatively high tenure in the field* considering their age. Also, the results of their responses suggest that they are *intrinsically motivated* to help others, not necessarily for enjoyment but for learning with the others. In addition, their answers indicate that *reputation plays a role for them*. However, as the individuals were rather unacquainted to each other, *centrality* in the network of participants must be assumed to be rather low. This on the other hand, is an aspect that this thesis' approach can help to improve.

What's more, the survey results suggest that *contextual overlap results in decreased effort* for message formulation. It hence may *contribute to more messages* as indicated by more than a third of the respondents who would respond more often when knowing that the requester is in a similar context. Also the results showed that being able to *post anonymously was an important feature* while at the same time generalized *reciprocity was relevant to the participants*. Taking into account that there were high levels of trust in this social fabric as stated by the participants they were motivated to contribute their knowledge within the implementation of this thesis' approach.

8 Conclusion and Outlook

This thesis described an approach for *fostering knowledge transfer* between individuals. For this purpose the *overlap of relevant activities* is employed as mechanism to *determine suitable interaction partners*. Relevant activities are deduced from IT-supported tasks that indicate *level of expertise* on the one hand and (work) *contextual similarity* between individuals on the other hand. The advantage of this thesis' approach is described in comparison to contemporary approaches – Yellow Page Systems, Expert Recommender Systems and three different kinds of Knowledge Networks. These contemporary approaches were also used to derive fundamental requirements for this thesis' approach. Additionally, influencing factors that determine the success of KM initiatives in general and KM systems in particular were taken into account to design an artifact that helps in fostering knowledge transfer. Next to a framework that supports the implementation of this thesis' approach, the approach itself was instantiated and evaluated in two case studies. The results of applying the proposed approach in the two case studies within an SAP system landscape could support its benefits in real-life usage.

8.1 Summary of Results

This thesis' progress was guided by four research questions. Consequently, the presentation of the summary of results will be arranged in accordance with them.

8.1.1 Research Question 1

The first research question focused on the elicitation of aspects that have to be considered when constructing a new KM approach. This was necessary as this thesis' overarching goal is to better support knowledge transfer – a central aspect of knowledge management. The first research question therefore asked:

Research Question 1 ***Which requirements have to be taken into account for the design of an IT-supported person-to-person Knowledge Management approach?***

The goal to determine relevant requirements that this thesis' approach needs to fulfill, was approached on different paths. First, advantages and disadvantages of contemporary approaches that support knowledge transfer were determined as those provide well-proven, real-life insight from a practical point of view. For this purpose, the three common approaches *Yellow Page Systems (YPS)*, *Expert Recommender Systems (ERS)* and *Knowledge Networks* were investigated. For each of them diverse challenges could be determined. Altogether five areas for improvement for YPS and another five for ERS could be identified. Knowledge network is a collective term for related approaches which is why the subtypes *Community of*

Practice, *Distributed Network of Practice* and *Electronic Network of Practice* were separately analyzed. For each of these types of knowledge networks seven challenges could be identified. Using all challenges, a set of conclusions was derived that informs the design of a more suitable approach. These conclusions served as a first source of requirements for this thesis' new approach.

In a next step, influencing factors were determined that need to be taken into account when implementing knowledge management approaches. Hence, requirements on a more abstract level were analyzed as they apply for KM approaches in general. Performing a literature review, twelve critical success factors could be identified that influence the result of a KM initiative, which were arranged according to the three dimensions *organization*, *person*, and *technology*. The results of this analysis informed the later design and served as second source of requirements. Additionally, a literature review focusing on IT-based knowledge management systems was performed. Since this thesis' approach will be IT-based the second literature review provided further insight in requirements that need to be addressed. In the course of this analysis eleven critical success factors for the IT part of IT-based knowledge management initiatives were extracted.

The output of *Design Science Research* includes additions to the relevant knowledge base (see subchapter 1.4). In light of the first research question those additions to the knowledge base are as follows. The analysis of the three common approaches that support direct knowledge transfer – YPS, ERS and Knowledge Networks – allows a deeper understanding of challenges and advantages that come with the implementation of them. The list of challenges for each approach may serve as a starting point for improving the approaches or for implementing a new approach as done in this thesis. After the design of the approach or the change to the existing one, the list of challenges acts as a reference for checking on the achieved improvement. Hence, this thesis methodologically extends the knowledge base by providing *validation criteria* (Hevner et al. 2004) for approaches that support direct knowledge transfer. Further, the analysis of critical success factors for KM initiatives aggregates the results of previous research on KM success. It harmonizes the terminology and hence offers additions to the knowledge base in the form of those harmonized and defined *constructs* (Hevner et al. 2004) that influence the success of KM initiatives. Additionally, the constructs are arranged according to a well-accepted schema to provide the knowledge base with a first, rough *model* (Hevner et al. 2004) of interrelations between them. The same applies to the critical success factors for IT-based KMS. Here also the prevailing understanding was harmonized to derive a set of common *concepts* that were arrange in a first *model* that illustrates the interrelations between them.

8.1.2 Research Question 2

The second research question focused on the design of this thesis' KM approach. It adopts the *overlap of IT-supported tasks* as its core concept and this way addresses all the requirements that were derived in the previous research question. The second research question therefore asked:

> ***Research Question 2 What are the necessary elements of a Knowledge Management approach that builds upon the overlap of IT-supported tasks?***

To properly design this thesis' approach, first, design considerations were described and the selection among those choices was explained. This thesis' approach essentially is a combination and further improvement of *Knowledge Networks* and *ERS*. To combine the two the overlap of *IT-supported tasks as binding element* is utilized. This thesis' interpretation of IT-supported tasks is given and also reason is provided why this selection is senseful. Then a description is given of how the challenges of the influencing approaches such as *embeddedness* into operational tasks and systems, *information overload* and the treatment of *anonymity* form design dimensions for this thesis' approach. In accordance with the concept of *theory-driven design* (Card 1989; Briggs 2006), subsequently theoretical models were investigated to further guide the design process. To this respect, models that describe interpersonal knowledge transfer were analyzed and described, that rely on *Social Capital Theory* and *Problem Solving Theory*.

Based on this fundament, the design of this thesis' KM approach was derived. It features three layers: the *Representation Layer*, the *Mediation Layer* and the *Data Acquisition Layer*. The *Representation Layer* has substantial similarity with Knowledge Networks and acts as the frontend to the knowledge worker. It provides three categories of functionalities: *Network Size Management*, general *Management Aspects of the Network* and *User Interaction.* Functionalities for Network Size Management allow creating networks with the right number of participants, by offering ways to initiate networks, merge, or split them. The management of the network is supported in various ways. A mechanism allows to spread dyadic learning effects to a larger group of stakeholders. Further, individuals may assume different roles in the networks. In addition, different levels of anonymity are possible and the effect of each of them on this thesis' approach is emphasized. Finally, incentive mechanisms are part of this category. User Interaction is a fundamental concept in this thesis' approach. The user context is taken into account and details are provided about how the interactions with operational IS can be tailored. The structuring of interactions within the network is also part of this category, as are measures to increase the likelihood of contributions.

In the *Mediation Layer*, the *context and expertise models* of the individuals are used to create context-dependent connections between them when an individual seeks knowledge exchange with other individuals. The importance of both, *context overlap and expertise*, is emphasized and reasoned for when fostering the transfer of knowledge between individuals. This in turn motivated the use of *task overlap* as core principle. Showing possibilities to determine overlap the concept of *similarity* is formalized and different kinds of similarity measures were differentiated. As part of the design of the *Mediation Layer* the challenge of limiting the consumption of collective attention (Ye et al. 2008) was addressed as well. The way in which this thesis' approach can contribute to a limitation of wasted attention is therefore detailed.

The "lowest" level of this thesis' approach, the *Data Acquisition Layer,* concerns itself with acquiring the necessary information to build context and expertise models. Here, the question, which source systems could prove to be usable for this thesis' approach, was addressed. In addition, the question which data entities are necessary to support this thesis approach was tackled. The class of *Process-Aware Information Systems,* which comprises every IS that has an explicit or implicit process model, turned out to be viable sources. The data entities need to contain information about what has been done (*task*), who has done the task (the *knowledge workers* of interest), when the task has been done (for recreation of *sequences*) and possibly *payload data* that is specific for the application. Further, the Data Acquisition Layer includes design choices with respect to architectural considerations. As different IS may be relevant for the determination of expertise and context, a single system should integrate the information of different IS, while ad-hoc access to each single IS must be possible at the same time to have up-to-date user models.

The output of the second research question's results are twofold. For the affected environment, i.e. organizations that strive to foster knowledge transfer among their employees, this thesis' approach offers a *blueprint* to implement an IT-based KMS that supports their efforts. As this thesis' approach is generic by design, the actual implementation and some design choices within the approach need to be adapted to the contextual setting of the company. One such instantiation was developed and tested in two case studies in the course of research question four. On the other hand, this thesis' approach extends the knowledge base. Hevner et al. state that "[…] design-science research addresses important unsolved problems in unique or innovative ways or solves problems in more effective or efficient ways" (Hevner et al. 2004, p. 81) and that it is different from mere routine design when there is a "[…] clear identification of a contribution to the archival knowledge base of foundations and methodologies." (Hevner et al. 2004, p. 81). This thesis approached the challenge of fostering knowledge transfer, which has been addressed before, but did so using an *innovative way* by relying on task overlap as its central concept. Backed by further analysis of influencing factors and derivation from theoretical models, the challenges that contemporary approaches face were addressed, which renders this thesis' approach *more effective and efficient*. This thesis' approach is generic by design. This way it is applicable to a class of problems and presents an addition to the *foundations* of the knowledge base, as it may be reused and adapted to specific settings.

8.1.3 Research Question 3

In the third research question this thesis addressed possibilities to support the approach from a technological point of view. This thesis' approach is generic and adapts to the needs of its fields of application which is why many aspects are specific to the organizational context when seen from an implementation point of view. However, the central aspect, i.e. the *determination of the overlap of IT-supported tasks*, may be supported by a ready-made framework. The third research question therefore asked:

Research Question 3 ***How can the core aspect of this thesis' Knowledge Management approach - the determination of the overlap of IT-supported tasks - be supported?***

To derive a framework that provides utility for the determination of a suitable measure to determine overlap, as a first step requirements for it were gathered. It emerged that the determination of overlap, or similarity of instances respectively, essentially consists of the three steps: *(raw) data acquisition*, *application of similarity measure*, and *check of fitness* of the result. Hence, the requirements were elicited for each of these stages. As a result six requirements for modules that support data acquisition in this scenario, seven requirements for modules that support similarity determination and another one for modules that check the fitness of the result were logically deduced. Subsequently, the supportive framework consisting of five modules was designed. The first module, the *import modules*, handles the acquisition of raw data, while the second module, the *classification module*, allows to extend the data with classifications that may be used for machine learning purposes when "learning" a suitable measure. The third module, the *mining module*, deals with the (re-)creation of process models from logged instances – also known as *process mining* – which enables the use of structural similarity measures that otherwise would not be possible. The fourth module, the *similarity module*, includes functionality to determine numeric values that represent similarity using various (combinations of) algorithms. The final module, the *application module*, represents the target application or the relevant part of it that consumes the result and indicates the fitness of the similarity results. To support a broad range of possible input sources, the framework employs a very generic data format. Additionally, it includes a newly designed mechanism to enable the transformation of data stored in "snapshots" into data stored as events.

The main contribution of this research questions lies in the supportive framework for similarity measure determination. The knowledge base is extended with a ready-to-use *framework* (Hevner et al. 2004) that finds a suitable measure to determine the overlap of tasks with respect to a utility function. In the case of this thesis, the utility function is the appropriateness of interaction partners. Practitioners on the other hand are relieved of a tedious task, as the determination of similarity measures that fit a context is a computationally intensive task that also asks for a lot of experience on the implementer's side. Additionally, this thesis contributes a *data transformation mechanism* to the knowledge base. The mechanism was formalized as an XML dialect. It allows to transform two data storage paradigms into one another: one that stores *data snapshots* and one that stores *changes to data states* respectively. This enriches the knowledge base's *methodologies* with a *formalism* (Hevner et al. 2004) for data storage transformation.

8.1.4 Research Question 4

The fourth and last research question investigated the real-life applicability of this thesis' approach and shed light on the possible benefits that it may generate. The fourth research question therefore asked:

Research Question 4 ***What effects can be achived using this thesis' Knowledge Management approach and which implications for further research and development can be deduced?***

The evaluation of this thesis' approach was done with respect to different aspects. First, a *descriptive analysis* was performed in which the features of this thesis' approach were compared with the ones of contemporary approaches. Logical deduction and argumentative reasoning showed that this thesis' approach provides comparable advantages with respect to the other investigated approaches but is less or not at all affected by their disadvantages. As a next step in the evaluation of this thesis' approach, an *architecture analysis* was performed. This way the technological feasibility of this thesis' approach could be verified. It was possible to implement this thesis' approach, at first focusing on its data extraction capabilities, into an IT landscape with more than 100 SAP systems. Up to the day of writing, the implementation could collect some 2 million tasks that were performed by users in the SAP landscape starting July 2009. The process is automated and puts only minimal burden on the operational systems.

Then, the approach was evaluated in a real-life setting in two iterations. In the first iteration, this thesis' approach was implemented in an SAP system that was used by a small-scale group over several months. The qualitative feedback that could be collected from this group gave valuable insight into the applicability of this thesis' approach. It suggested that the very idea of this thesis' approach was well received, but that some implementation aspects could be improved. Additionally, the feedback allowed to critically judge the applicability of this thesis' approach in dependence of a group's characteristics. It turned out that, especially in networks that capitalize on strong ties, where individuals know each other well already, established means to transfer knowledge are likely preferred over using this thesis' approach.

In the second iteration of real-life evaluation, therefore, a larger-scale group of almost 100 individuals was addressed in group members did not know each other at all or only to a very limited degree. To enhance the user experience, the implementation-specific suggestions for improvement that were part of the first case study's feedback were included in a refined implementation. The participants in the second case study again used a SAP system for their work, but had different tasks and, this time, worked alone and not in teams as the first group of participants did. The results of this case study were acquired through different means. After the participants had finished their tasks, an online survey was distributed to them in which different constructs asked for the participants opinion about some of this thesis' design considerations. In addition, some constructs were included that evaluated this thesis' approach and its implementation with respect to the users' perceived value. Next to the survey data, the approach's implementation was evaluated using a click stream analysis to better understand real usage. The results indicated a high degree of user acceptance for this thesis' approach and also high perceived value.

The results of this research question not only show the benefits and challenges of this thesis' approach but also contribute to the relevant knowledge base. This thesis' approach as described in the course of the second research question was generic. In this research question an *instantiation* (Hevner et al. 2004) of it was created that, in the sense of Design Science Research, is part of the foundations of the knowledge base. It proves real-life applicability, shows advantages and challenges and yields empirical results of the application of this thesis' approach.

8.2 Limitations

The evaluation of this thesis' approach could show that it has advantages in comparison to contemporary approaches, it is implementable and the instantiation of the approach that was used in two case studies delivered obvious benefit to its stakeholders. However, these results have to been seen in light of the limitations of the evaluation.

The advantages that were worked out in comparison to contemporary approaches were deduced logically taking into account the features of each approach. While the deduction is senseful, support by empirical evidence would underline the superiority of this thesis' approach. Specifically, it would make sense to compare the usage and perceived benefit of the different approaches, for example in controlled experiments.

The implementability of this thesis' approach was proven in one large landscape that consisted of SAP systems. While the requirements that this thesis' approach imposes on the data are small and allow to generalize the applicability to all kinds of Process-Aware Information

Systems from a technical point of view, the same may not be true for the relative effort of implementation and maintenance. It is possible that for other kinds of IS the implementation and maintenance costs outweigh the benefits generated by the implementation of this thesis' approach. However, most sophisticated IS offer suitable interfaces to interact with custom implementations. Therefore, it is possible that the effort outweighs the benefit but this is rather unlikely.

The case studies showed significant benefit for the stakeholders. However, the case studies were not within a corporate setting. Small teams of already acquainted individuals were used and additionally a larger set of unacquainted individuals that had to perform knowledge-intensive tasks and at the same time learned the usage of the system. While this is similar to newcomers in a corporation, the general transferability of the results cannot be guaranteed but only be reasoned to be likely. Yet it might be the case that the perceived benefit may be different for corporate users. In addition, the number of participants was limited in the case studies to roughly twenty in the first case study and about 100 in the second case study. One can only project but not justify that the results that were obtained transfer to larger settings. However, the larger the groups the higher the chance of individuals having contextual and temporal overlap and hence the higher the chance that individuals contribute and as a consequence the positive effect of this thesis' approach should arguably be larger not smaller. The limited size of the case study groups also limited the functionality of this thesis' approach that were actually implemented in the prototypes. For example, functionalities for Network Size Management were not an issue in those small networks created in the case study which is why functionality for managing the size was not implemented. Consequently this aspect of this thesis' approach could not be validated.

Furthermore, in corporate settings, another aspects becomes important, that was less so in the case studies. Depending on the country in which this thesis' approach is implemented, there are laws that protect personal information and disallow the implementation of facilities that are capable of monitoring person-specific work behavior. This thesis' approach has no intention to invade personal privacy and is not meant to monitor the work behavior of individuals for performance evaluation. However, it could be used this way, which is why it falls under the law's scope and hence may be limited in its ability to support knowledge transfer due to legal reasons.

Some results of the feedback obtained from the case studies challenge some of this thesis' assumptions and hence limit the impact of the approach. In the design of this thesis' approach the effects of other, already established means of interaction were not catered for. In the case studies it turned out that those other "communication channels" and this thesis' approach interfere with each other in terms of usage and perceived value. In addition, the results of the second case study suggest that the usage of suitable measures for determining expertise is a complex issue. The survey results indicated that not all participants felt as experts while the tool deemed them to be one. This is natural as in the implementation, the contextual overlap between two individuals was on purpose overemphasized in comparison to evidence for expertise to counterbalance a lack in temporal overlap of the individual's work contexts. This turned out to be a good choice as the survey results suggest that contextual overlap was more important to the participants. However, it sheds light on a limitation of this thesis' approach. While this thesis additionally provides a framework for determining suitable algorithms in a machine learning-based approach, it is still a difficult endeavor to find a proper balance between level of expertise and contextual overlap. This challenge of properly balancing those two aspects is left to the organization implementing the approach and likely requires some experimentation and hence shows a partial fuzziness of this thesis' approach that is inherent to the hard-to-formalize nature of human interaction.

8.3 Outlook

In the course of this thesis the initial goals could be fulfilled. Inspired by the increasing importance of knowledge transfer processes in organizations and the identified deficiencies of contemporary approaches, this thesis set out to develop an approach that, by building upon innovative methods, would overcome the current limitations. It was possible to show how this thesis' approach that is based on *task context overlap* has advantages compared to the contemporary approaches. Also it was possible to show how it delivers value to its stakeholders in the two case studies that were performed. However, the case studies and the limitations of this thesis' approach that were worked out in the previous sub-chapter fuel the need for more research into this thesis' approach and its application. In the following, those research issues are addressed for which large potential for valuable contribution is expected. Also, first directions for the further research in these areas is provided.

To obtain a better empirical basis on which future research could build on, it would be a valuable endeavor to employ this thesis' approach in more and different settings. Especially, it would be interesting to see its application within a group of stakeholders that, due to a larger number of participants, has a large chance of two individuals actually working on similar or the same tasks at the very same moment. The possibility of immediate interaction is likely to trigger different results, as the interaction analysis of the second case study also suggests. There, it was also the case that interactions were more likely whenever usage of the application was high also. Extensions to this thesis' approach and its applicability could also be informed by cross-cultural application of instantiations of the approach. This is an aspect that was not adopted in the initial design, but that very well could influence the success of it when applied in different cultural settings, especially concerning appreciation of the contribution of others and effects of anonymity that are rooted deeply in cultural imprinting.

The second case study showed that while the ones that did use the approach appreciated it, the overall use of the support offered by the implementation was only moderate among the participants. Given the high satisfaction and perceived value as determined by the survey results, this raises the question how a wider use of this thesis' approach could be achieved. Additionally, the second case study surfaced another interesting field of study that is related to the previous question. There was another established communication channel, an internal forum, that the participants had been using also before being exposed to this thesis' supportive tool and afterwards as well. The implementation of this thesis' approach and the established communication channel seemed to cannibalize on each other, which is why an interesting endeavor for further research would be to understand the interactions between this thesis' approach and others that are established already and which factors influence the preference for the one or the other.

Finally, at its core, this thesis' approach relies on the expertise of one individual and at the same time on the contextual overlap of the requesting and the helping individual. While both aspects naturally contribute to successfully establishing a connection between knowledge seeker and knowledge bearer, it is unknown which contributes more, how those two aspects interact exactly and which external factors put more favor on the one side than on the other. Future research analyzing the contribution behavior of knowledge workers in regard to this would substantiate the workings of this thesis' approach.

List of References

Aalst, W. M. van der;Beer, H. T. de;Dongen, B. F. van (**2005**): Process mining and verification of properties: An approach based on temporal logic. In R. Meersman, Z. Tari (Eds.): On the move to meaningful internet systems 2005. Proceedings of CoopIS, DOA, and ODBASE: OTM Confederated International Conferences CoopIS, DOA, and ODBASE 2005. Agia Napa, Cyprus, October 31-November 4, 2005: Springer, pp. 130–147.

Aalst, W. van der;Weijters, T.;Maruster, L. (**2004a**): Workflow mining: Discovering process models from event logs. In *Knowledge and Data Engineering, IEEE Transactions on* 16 (9), pp. 1128–1142.

Aalst, W. van der;Medeiros, A. K. A. de (**2005a**): Process Mining and Security: Detecting Anomalous Process Executions and Checking Process Conformance. In *Electronic Notes in Theoretical Computer Science* 121, pp. 3-21.

Aalst, W. van der;Weijters, A. J. M. M.;Maruster, L. (**2004b**): Workflow Mining: Discovering Process Models from Event Logs. In *IEEE transactions on knowledge and data engineering* 16 (9), pp. 1128-1142.

Aalst, W. van der;Medeiros, A. de;Weijters, A.: Genetic Process Mining, pp. 48–69.

Aalst, W. van der;Dongen, B. F. van;Herbst, J. et al. (**2003**): Workflow Mining: A Survey of Issues and Approaches. In *Data and Knowledge Engineering* 47 (2), pp. 237-267.

Aalst, W. van der;Weijters, A. J. M. M.;Maruster, L. (**2004c**): Workflow Mining: Discovering Process Models from Event Logs. In *IEEE transactions on knowledge and data engineering* 16 (9), pp. 1128-1142.

Aalst, W. van der;Medeiros, A. K. A. de (**2005b**): Process Mining and Security: Detecting Anomalous Process Executions and Checking Process Conformance. In *Electronic Notes in Theoretical Computer Science* 121, pp. 3-21.

Abecker, Andreas; Hinkelmann, Knut; Maus, Heiko; Müller, Hans Jürgen (**2002**): Geschäftsprozessorientiertes Wissensmanagement effektive Wissensnutzung bei der Planung und Umsetzung von Geschäftsprozessen. 1st. Berlin: Springer (Xpert.press). Available online at "http://www.bsz-bw.de/rekla/show.php?mode=source&eid=UNI_0_9797612inh Inhaltsverzeichnis".

Accenture (**2005**): IT Investing for High Performance: A Global Survey of CIOs. Available online at http://www.accenture.com/Global/Research_and_Insights/By_Role/HighPerformance_IT/CIOResearch/ITInvestHPSurveyCIOs.htm.

Ackerman, M. S.;Halverson, C. (**2004**): Sharing expertise: The next step for knowledge management. In Marleen Huysman, Volker Wulf (Eds.): Social capital and information technology. Cambridge, Mass: MIT Press.

Adomavicius, G.;Tuzhilin, A. (**2005**): Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. In *IEEE transactions on knowledge and data engineering* 17, pp. 734–749.

Agrawal, R.;Gunopulos, D.;Leymann, F. (**1998**): Mining Process Models from Workflow Logs. In Hans-Jörg Schek, Fèlix Saltor and Isidro Ramos and Gustavo Alonso (Eds.): Proceedings of Advances in Database Technology - EDBT‘98, 6th International Conference on Extending Database Technology. International Conference on Extending Database Technology. Valencia, Spain,, March 23-27, 1998: Springer (LNCS, 1377), pp. 469–483.

Ahmed, P. K.;Lim, K. K.;Zairi, M. (**1999**): Measurement practice for knowledge management. In *Journal of Workplace Learning* 11 (8), pp. 304–311.

Alavi, M.;Leidner, D. E. (**2001**): Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. In *MIS Quarterly* 25, pp. 107–136. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=4386667&site=ehost-live".

Al-Mabrouk, K. (**2006**): Critical Success Factors Affecting Knowledge Management Adoption: A Review of the Literature. In *Innovations in Information Technology, 2006*, pp. 1–6.

APQC (**2000**): Executive Summary - Successfully Implementing Knowledge Management. Houston, Texas: American Productivity & Quality Center, pp. 1–15.

Archer, N. (**2006**): Classification of Communities of Practice. In E. Coakes, S. Clarke (Eds.): Encyclopedia of communities of practice in information and knowledge management: IGI Global, pp. 21–29.

Armstrong, David Malet (**2008**): Belief truth and knowledge. Repr. Cambridge: Cambridge Univ. Press.

Arora, R. (**2002**): Implementing KM – a balanced score card approach. In *Journal of Knowledge Management* 6 (3), pp. 240–249.

Augustin, Siegfried (**1990**): Information als Wettbewerbsfaktor. Informationslogistik - Herausforderung an das Management. Zürich [u.a.]: Verl. Industrielle Organisation [u.a.] (Schriftenreihe Erfolgspotentiale für Unternehmer und Führungskräfte).

Baker, M. J. (**2000**): Writing a Literature Review. In *Marketing Review* 1 (2), pp. 219–247. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=4389483&site=ehost-live".

Balabanović, M.;Shoham, Y. (**1997**): Fab: content-based, collaborative recommendation. In *Commun. ACM* 40, pp. 66-72. Available online at http://doi.acm.org/10.1145/245108.245124.

Balog, K.;Rijke, M. de (**2006**): Finding experts and their details in e-mail corpora. In Arun Iyengar, David C. de Roure, Leslie Carr (Eds.): Proceedings of the 15th International Conference on World Wide Web. Edinburgh, Scotland,, May 23-26, 2006. New York, N.Y: ACM.

Balog, K.;Rijke, M. de (**2007**): Finding similar experts. In : SIGIR '07 Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. Association for Computing Machinery. New York, NY: ACM (Special issue of the SIGIR forum). Available online at http://doi.acm.org/10.1145/1277741.1277926.

Barnes, Stuart J. (**2002**): Knowledge management systems theory and practice. 1st. London: Thomason Learning. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=009539888&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Becerra-Fernandez, I. (**2000**): Facilitating the online search of experts at NASA using expert seeker people-finder. Proceedings of the 3rd International Conference on Practical Aspects of Knowledge Management (PAKM), Basel, Switzerland. In Ulrich Reimer (Ed.): Proceedings of the Third International Conference on Practical Aspects of Knowledge Management. PAKM 2000. Basel, Switzerland, October 30-31, 2000: CEUR-WS.org (CEUR Workshop Proceedings, 34).

Becerra-Fernandez, I. (**2001**): Searching for Experts with Expertise-Locator Knowledge Management Systems. Edited by Florida International University.

Becker, J. (**2010**): Prozess der gestaltungsorientierten Wirtschaftsinformatik. In Hubert Österle, Robert Winter, Walter Brenner (Eds.): Gestaltungsorientierte Wirtschaftsinformatik. Ein Plädoyer für Rigor und Relevanz. Nürnberg: Infowerk, pp. 13–17.

Becker, Jörg; Krcmar, Helmut; Niehaves, Björn (Eds.) (**2009**): Wissenschaftstheorie und gestaltungsorientierte Wirtschaftsinformatik. Heidelberg: Physica-Verl. Available online at http://www.gbv.de/dms/zbw/599825308.pdf.

Beckermann, A. (**2011)**: Zur Inkohärenz und Irrelevanz des Wissensbegriffs. Plädoyer für eine neue Agenda in der Erkenntnistheorie. In *Zeitschrift für philosophische Forschung* 55, pp. 571–593.

Beham, G.;Kump, B.;Ley, T. et al. (**2010)**: Recommending knowledgeable people in a work-integrated learning system. In *Procedia Computer Science* 1 (2), pp. 2783–2792. Available online at http://www.sciencedirect.com/science/article/B9865-50YNHC8-3/2/0d02c87d3ee51a3e1070d7c8964faf45.

Bellmann, Matthias; Krcmar, Helmut; Sommerlatte, Tom (Eds.) (**2002)**: Praxishandbuch Wissensmanagement Strategien - Methoden - Fallbeispiele. 1st. Düsseldorf: Symposion Publ. (Wissensmanagement). Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=010192154&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Bender, Thomas (**1982)**: Community and social change in America. Baltimore: Johns Hopkins Univ.Pr (A Johns Hopkins paperback).

Bettoni, M.;Clases, C.;Wehner, T. (**2004)**: Communities of Practice im Wissensmanagement: Charakteristika, Initiierung und Gestaltung. In Gabi Reinmann, Heinz Mandl (Eds.): Psychologie des Wissensmanagements. Perspektiven, Theorien und Methoden. 1.th ed. Göttingen: Hogrefe, pp. 319-226.

Billsus, D.;Brunk, C. A.;Evans, C. et al. (**2002)**: Adaptive interfaces for ubiquitous web access. In *Commun. ACM* 45, pp. 34-38. Available online at http://doi.acm.org/10.1145/506218.506240.

Binney, D. (**2001)**: The knowledge management spectrum - understanding the KM landscape. In *Journal of Knowledge Management* 5 (1), pp. 33–42.

BITKOM (**2007)**: Wichtige Trends im Wissensmanagement 2007 bis 2011 - Positionspapier des BITKOM. Available online at http://www.bitkom.org/files/documents/Trendreport_WM_zur_KnowTech2007.pdf.

Blau, P.M (**1986)**: Exchange and power in social life: Transaction Publishers.

Blessing, D. (**2001)**: Content Management für das Business Engineering. Dissertation, St. Gallen. Hochschule für Wirtschafts-, Rechts und Sozialwissenschaften (HSG).

Blohm, I.;Ott, F.;Huber, M. et al. (**2010)**: Extending Open Innovation Platforms into the real world – Using Large Displays in Public Spaces. In : European Academy of Management EURAM, vol. 2010: Eigenverlag.

Böhmann, T.;Krcmar, H. (**2002)**: Werkzeuge für das Wissensmanagement. In *M. Bellmann, H. Krcmar, T. Sommerlatte, Wissensmanagement, Symposion Publishing GmbH, D�sseldorf.*

Braganza, A.;Möllenkramer, G. J. (**2002)**: Anatomy of a failed knowledge management initiative: lessons from PharmaCorp's experiences. In *Knowl. Process Mgmt* 9 (1), pp. 23–33.

Briggs, R. (**2006)**: On theory-driven design and deployment of collaboration systems. In *International Journal of Human-Computer Studies* 64 (7), pp. 573–582.

Bromme, R.;Rambow, R.;Sträßer, R. (**1996)**: Jenseits von ‚Oberfläche'und ‚Tiefe': Zum Zusammenhang von Problemkategorisierungen und Arbeitskontext bei Fachleuten des Technischen Zeichnens. In *Expertiseforschung: Theoretische und methodische Grundlagen*, pp. 150–168.

Bromme, R.;Jucks, R.;Rambow, R. (**2004)**: Experten-Laien-Kommunikation im Wissensmanagement. In Gabi Reinmann, Heinz Mandl (Eds.): Psychologie des Wissensmanagements. Perspektiven, Theorien und Methoden. 1.th ed. Göttingen: Hogrefe, pp. 176–188.

Brown, J. S.;Duguid, P. (**1991)**: Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. In *Organization science* 2 (1), pp. 40–57.

Brown, J.;Duguid, P. (2001): Knowledge and organization: A social-practice perspective. In *Organization science* 12 (2), pp. 198–213.

Brügge, Bernd; Dutoit, Allen H. (2004): Object-oriented software egineering using UML, patterns, and Java. 2nd. Upper Saddle River, NJ: Pearson Prentice Hall (An Alan R. Apt book).

Bullen, Christine V.; Rockart, John F. (1986): The rise of managerial computing the best of the Center for Information Systems Research, Sloan School of Management, Massachusetts Institute of Technology. 1st. Homewood, Ill.: Dow Jones-Irwin.

Bullinger, Hans-Jörg; Wildemann, Horst (2002): Wissensmanagement Wissen als strategische Ressource im Unternehmen. 2nd. München: TCW Transfer-Centrum (TCW-Report; 30).

Bullinger, Hans-Jörg; Wörner, Kai; Prieto, Juan (1997): Wissensmanagement heute Daten, Fakten, Trends. 1st. Stuttgart: Fraunhofer-Inst. für Arbeitswirtschaft und Organisation (IAO). Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=008403821&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Burke, R. (2000): Knowledge-Based Recommender Systems. In Allen Kent (Ed.): Encyclopedia of library and information science. 69th ed. New York: Dekker (32).

Burke, R. D.;Hammond, K. J.;Young, B. C. (1997): The FindMe Approach to Assisted Browsing. In *IEEE Expert* 12 (4), pp. 32-30. Available online at http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.88.5033.

Byrne, D.E (1971): The attraction paradigm: Academic Press.

Campbell, C. S.;Maglio, P. P.;Cozzi, A. et al. (2003): Expertise identification using email communications. Proceedings of the twelfth international conference on Information and knowledge management. In Ophir Frieder (Ed.): Proceedings of the Twelft International Conference on Information and Knowledge Management. International Conference on Information and Knowledge Management; CIKM 2003. New Orleand, Louisiana, USA, November 3 - 8, 2003. ACM. New York, NY: ACM Press.

Card, S. (1989): Theory-driven design research. In *Applications of Human Performance Models to System Design*, pp. 501–509.

Chase, R. L. (1997): The Knowledge-Based Organization: An International Survey. In *Journal of Knowledge Management* 1 (1), pp. 38–49.

Chourides, P.;Longbottom, D.;Murphy, W. (2003): Excellence in knowledge management: an empirical study to identify critical factors and performance measures. In *Measuring Business Excellence* 7 (2), pp. 29–45.

Chua, A.;Lam, W. (2005): Why KM projects fail: a multi-case analysis. In *Journal of Knowledge Management* 9 (3), pp. 6–17.

Clark, Herbert H. (2007): Using language. 7. print. Cambridge: Cambridge Univ. Press.

Constant, D.;Sproull, L.;Kiesler, S. (1996): The kindness of strangers: The usefulness of electronic weak ties for technical advice. In *Organization science* 7 (2), pp. 119–135.

Craswell, N.;Hawking, D.;Am Vercoustre et al. (2001): P@ noptic expert: Searching for experts not just for documents. Ausweb Poster Proceedings, Queensland, Australia. In : Ausweb Poster Proceedings. Queensland, Australia.

Daft, R.;Lengel, R. (1984): Information Richness. A new approach to managerial behavior and organization design. In L.L Cummings, B.M Staw (Eds.): Research in organizational behavior 6. Homewood, Ill.: JAI Press, pp. 191–233.

Daniel, Alexander (2001): Implementierungsmanagement. Ein anwendungsorientierter Gestaltungsansatz. 1st ed. Wiesbaden: Dt. Univ.-Verl.

Daniel, D. R. (**1961**): Management Information Crisis. In *Harvard Business Review* 39 (5), pp. 111–121. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=6780970&site=ehost-live".

Darlington, Keith (**2000**): The essence of expert systems. Harlow: Prentice Hall (The essence of computing series).

Davenport, E.;Hall, H. (**2002**): Organizational knowledge and communities of practice. In *Annual review of information science and technology* 36 (1), pp. 170–227.

Davenport, T. H. (**1997**): Ten principles of knowledge management and four case studies. In *Knowledge and Process Management* 4 (3), pp. 187–208.

Davenport, T. H. (**2011**): Rethinking knowledge work: A strategic approach. Edited by McKinsey Quarterly. McKinsey & Company. Available online at https://www.mckinseyquarterly.com/Organization/Strategic_Organization/Rethinking_knowledge_work_A_strategic_approach_2739.

Davenport, T. H.;Völpel, S. C. (**2001**): The rise of knowledge towards attention management. In *Journal of Knowledge Management* 5 (3), pp. 212–222.

Davenport, Thomas H.; Prusak, Laurence (**2000a**): Working knowledge. How organizations manage what they know. Boston, Mass: Harvard Business School Press.

Davenport, Thomas H.; Prusak, Laurence (**2000b**): Working knowledge how organizations manage what they know. 1st. Boston, Mass.: Harvard Business School Press. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=015622258&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Davenport, T. H.;Long, D. W. de;Beers, M. C. (**1998**): Successful Knowledge Management Projects. In *Sloan Management Review* 39 (2), pp. 43–57. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=159474&site=ehost-live".

Davenport, Thomas H.; Probst, Gilbert J. B. (**2002**): Knowledge management case book Siemens best practises. 2nd. Erlangen: Publicis Corporate Publ.

Davis, F. D. (**1986**): A technology acceptance model for empirically testing new end-user information systems: Theory and results. Doctoral Dissertation. Massachusetts Institute of Technology. Sloan School of Management.

Davis, F. D. (**1989**): Perceived usefulness, perceived ease of use, and user acceptance of information technology. In *Mis Quarterly* 13 (3), pp. 319–340.

Deiters, W.;Lucas, R.;Weber, T. (**1999**): Skill-Management: a building block for project management with flexible teams. Dortmund (ISST Bericht).

Denzin, Norman K. (**1978**): The research act. A theoretical introduction to sociological methods. 2nd ed. New York: McGraw-Hill.

Dixon, Nancy M. (**2005**): Common knowledge how companies thrive by sharing what they know. 1st reprinted. Boston, Mass.: Harvard Business School Press.

Dongen, B. F. van;Medeiros, A. K. A. de;Verbeek, H. M. W. et al. (**2005**): The ProM Framework: A New Era in Process Mining Tool Support. In Gianfranco Ciardo, Philippe Darondeau (Eds.): Applications and Theory of Petri Nets 2005, 26th International Conference, ICATPN 2005, Miami, USA, June 20-25, 2005, Proceedings: Springer (Lecture notes in computer science, 3536), pp. 444–454.

Donoghue, L.;Harris, J.;Weitzman, B. (**1999**): Knowledge management strategies that create value. In *Accenture Outlook* Outlook (1), pp. 48–53. Available online at "http://www.accenture.com/NR/rdonlyres/9BC14B89-E2AC-4D5C-855D-23C8013FF4F6/0/knowledge2.pdf".

Drucker, P. (**1988**): The coming of the new organization. In *Harvard business review*.

Du Plessis, M. (**2007**): Knowledge management: what makes complex implementations successful? In *Journal of Knowledge Management* 11 (2), pp. 91–101.

Dumas, Marlon; van der Aalst, Wil; Ter Hofstede, Arthur (**2005**): Process-aware information systems. Bridging people and software through process technology. Hoboken, N.J: Wiley-Interscience. Available online at http://catdir.loc.gov/catdir/enhancements/fy0621/2005001369-b.html.

Eppler, Martin J. (**2001**): Fallstudien zum Wissensmanagement: Lösungen aus der Praxis aufbereitet für die Aus- und Weiterbildung. St. Gallen: NetAcad. Press.

Erlikh, L. (**2002**): Leveraging legacy system dollars for e-business. In *IT Professional* 2 (3), pp. 17–23.

Etzioni, A. (**1996**): The responsive community: A communitarian perspective. In *American Sociological Review*, pp. 1–11.

Fahey, L.;Prusak, L. (**1998**): The Eleven Deadliest Sins of Knowledge Management. In *California Management Review* 40, pp. 265–276. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=738868&site=ehost-live".

Felfernig, A.;Friedrich, G.;Jannach, D. et al. (**2006**): An integrated environment for the development of knowledge-based recommender applications. In *International Journal of Electronic Commerce* 11 (2), pp. 11–34.

Festinger, L. (**1954**): A theory of social comparison processes. In *Human Relations, 7(2) 117-140.* 7 (2), pp. 117–140.

Fettke, P. (**2006**): State-of-the-Art des State-of-the-Art. In *Wirtschaftsinformatik* 48 (4), pp. 257–266.

Folger, Robert G.; Cropanzano, Russell (**1998**): Organizational justice and human resource management. London: Sage.

Foster, A.;Ford, N. (**2003**): Serendipity and information seeking: an empirical study. In *Journal of Documentation* 59 (3), pp. 321–340.

Frauenhofer IAO und Deutsche Bank (**1999**): Wettbewerbsfaktor Wissen. Leitfaden zum Wissensmanagement. Frankfurt am Main.

Gabler Wirtschaftslexikon (**2011**): Gabler Wirtschaftslexikon, Stichwort: Wissen. Available online at http://wirtschaftslexikon.gabler.de/Archiv/75634/wissen-v3.html, checked on 28/02/2011.

Gettier, E. (**1963**): Is justified true belief knowledge? In *Analysis* 23 (6), pp. 121–123.

Geyer, W.;Dugan, C. (**2010**): Inspired by the audience: a topic suggestion system for blog writers and readers. In Kori Inkpen Quinn, Carl Gutwin, John C. Tang (Eds.): Proceedings of the 2010 ACM conference on Computer supported cooperative work. CSCW 2010. Savannah, Georgia, USA, February 06 - 10, 2010. New York, NY, USA: ACM.

Gödel, K. (**1931**): Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. In *Monatshefte für Mathematik* 38 (1), pp. 173–198. Available online at http://dx.doi.org/10.1007/BF01700692.

Goh, S. C. (**2002**): Managing effective knowledge transfer: an integrative framework and some practice implications. In *Journal of Knowledge Management* 6 (1), pp. 23–30.

Gottschalk, P. (**2000**): Knowledge Management Systems: A Comparison of Law Firms and Consulting Firms. In *Informing Science* 3 (3), pp. 117–124.

Götz, Klaus; Schmid, Michael (**2004**): Theorien des Wissensmanagements. Frankfurt am Main: Lang.

Granitzer, M.;Granitzer, G.;Tochtermann, K. et al. (**2009**): Automating knowledge transfer and creation in knowledge intensive business processes. In : Proceedings of the Business Process Management Workshops. Berlin;Heidelberg: Springer.

Grant, R. M. (**1996)**: Toward a knowledge-based theory of the firm. In *Strategic Management Journal* 17, pp. 109–122.

Greco, G.;Guzzo, A.;Pontieri, L. et al. (**2004)**: Mining Expressive Process Models by Clustering Workflow Traces. In Honghua Dai, Ramakrishnan Srikant, Chengqi Zhang (Eds.): Advances in Knowledge Discovery and Data Mining 8th Pacific-Asia Conference. PAKDD 2004. Sydney, Australia, May 26-28, 2004. Berlin, Heidelberg: Springer Berlin Heidelberg (3056), pp. 52–62.

Greco, G.;Guzzo, A.;Pontieri, L. (**2005)**: Mining Hierarchies of Models: From Abstract Views to Concrete Specifications. In W.M. P. van der Aalst, Boualem Benatallah, Fabio Casati, Francisco Curbera (Eds.): Business Process Management 3rd International Conference, BPM 2005. BPM 2005. Nancy, France, September 5-8, 2005. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture notes in computer science, 3649), pp. 32–47.

Greer, J.;McCalla, G.;Collins, J. et al. (**1998a)**: Supporting Peer Help and Collaboration in Distributed Workplace Environments. In *International Journal of AI and Education* (9), pp. 159–177.

Greer, J.;McCalla, G.;Cooke, J. et al. (**1998b)**: The Intelligent Helpdesk: Supporting Peer-Help in a University Course. In Barry Goettl, Henry Halff, Carol Redfield, Valerie Shute (Eds.): Intelligent Tutoring Systems, vol. 1452: Springer Berlin / Heidelberg (Lecture notes in computer science), pp. 494–503. Available online at http://dx.doi.org/10.1007/3-540-68716-5_55.

Gronau, N.;Uslar, M. (**2004)**: Integrating Knowledge Management and Human Resources Via Skill Management. In *Journal of Universal Computer Science: Proceedings of I-Know '04*, pp. 135–142.

Gunther, C. W.;van der Aalst, W. M. (**2006)**: A generic import framework for process event logs. In *Applications and Theory of Petri Nets 2005, 26th International Conference, ICATPN 2005, Miami, USA, June 20-25, 2005, Proceedings* 4103, p. 81.

Guy, I.;Jacovi, M.;Perer, A. et al. (**2010)**: Same places, same things, same people?: mining user similarity on social media. In Kori Inkpen Quinn, Carl Gutwin, John C. Tang (Eds.): Proceedings of the 2010 ACM conference on Computer supported cooperative work. CSCW 2010. Savannah, Georgia, USA, February 06 - 10, 2010. New York, NY, USA: ACM. Available online at http://doi.acm.org/10.1145/1718918.1718928.

Hansen, M. T.;Nohria, N.;Tierney, T. (**1999)**: What's your Strategy for Managing Knowledge? In *Harvard Business Review* 77, pp. 106–116. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=1613197&site=ehost-live".

Hasanali, Farida (**2002)**: Critical Success Factors of Knowledge Management: Published by the American Productivity & Quality Center.

Hasanali, F.;APQC (**2004)**: Critical Success Factors of Knowlege Management. In Michael E. D. Koenig, T. Kanti Srikantaiah (Eds.): Knowledge management lessons learned what works and what doesn't. 1st. Medford, N.J.: American Society for Information Science and Technology (ASIST monograph series), pp. 55–71.

Hasler Roumois, Ursula (**2007)**: Studienbuch Wissensmanagement Grundlagen der Wissensarbeit in Wirtschafts-, Non-Profit- und Public-Organisationen. 1st. Zürich: Orell Füssli (UTB ; 2954 : Wirtschaftswissenschaften, Betriebswirtschaftslehre). Available online at "http://deposit.d-nb.de/cgi-bin/dokserv?id=2966643&prov=M&dok_var=1&dok_ext=htm Inhaltstext http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=015692965&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Hecht, S.;Schmidl, J.;Weckenmann, D. et al. (**2008**): Data Warehouse Design - Teaching How to Design a Data Warehouse with Respect to Business Requirements. In : Proceedings of the 18th Workshop on Information Technologies and Systems. WITS. Paris, France.

Heckhausen, Jutta (**2006**): Motivation und Handeln. Heidelberg: Springer Medizin Verlag.

Heeren, F.;Sihn, W. (**2001**): XPERTFINDER - Nachrichtenanalyse zur themenfeldbezogenen Expertensuche. In Norbert Gronau (Ed.): Wissensmanagement. Systeme - Anwendungen - Technologien. Aachen: Shaker (Berichte aus der Wirtschaftsinformatik), pp. 81–92.

Heisig, Peter (**2005**): Integration von Wissensmanagement in Geschäftsprozesse. 1st. Berlin: Produktionstechnisches Zentrum Berlin (PTZ) (Berichte aus dem Produktionstechnischen Zentrum Berlin).

Helm, R.;Meckl, R.;Sodeik, N. (**2007**): Systematisierung der Erfolgsfaktoren von Wissensmanagement auf Basis der bisherigen empirischen Forschung. In *Zeitschrift für Betriebswirtschaft* 77 (2), pp. 211–241. Available online at "http://dx.doi.org/10.1007/s11573-007-0017-4".

Hevner, A. R.;March, S. T.;Park, J. et al. (**2004**): Design science in information systems research. In *Mis Quarterly* 28 (1), pp. 75–105.

Hofstede, G. (**1980**): Culture and Organizations. In *International Studies of Management & Organization* 10 (4), pp. 15–41. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=5812910&site=ehost-live".

Holsapple, C. W.;Joshi, K. D. (**2000**): An investigation of factors that influence the management of knowledge in organizations. In *The Journal of Strategic Information Systems* 9 (2-3), pp. 235–261. Available online at "http://www.sciencedirect.com/science/article/B6VG3-41V3437-8/2/9d02cb20aeedea520a152c351062b389".

Holsapple, C.;Whinston, A. (**1987**): Knowledge-based organizations. In *The Information Society* 5 (2), pp. 77–90.

Homans, G. (**1958**): Social behavior as exchange. In *American journal of sociology* 63 (6), pp. 597–606.

Jensen, M. B.;Johnson, B.;Lorenz, E. et al. (**2007**): Forms of Knowledge and Modes of Innovation. In *Research Policy* 36.

Jiang, Z.;Wang, W.;Benbasat, I. (**2005**): Multimedia-based interactive advising technology for online consumer decision support. In *Communications of the ACM* 48 (9), pp. 92–98.

Kanfer, R. (**1990**): Motivation theory and industrial and organizational psychology.

Kankanhalli, A.;Bernard C. Y. Tan;Wei, K.-k. (**2005**): Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation. In *Mis Quarterly* 29, pp. 113-143.

Kautz, H.;Selman, B.;Shah, M. (**1997**): The Hidden Web. In *AI Magazine* 18, pp. 27-36.

Kazi, Abdul Samad; Wolf, Patricia (**2006**): Real-Life Knowledge Management: Lessons from the Field. 1st: KnowledgeBoard.

Keller, G.;Nüttgens, M.;Scheer, A.-W. (**1992**): Semantische Prozeßmodellierung auf der Grundlage „Ereignisgesteuerter Prozeßketten (EPK)“. Universität des Saarlandes, Saarbrücken. Instituts für Wirtschaftsinformatik (IWi).

Kirkham, R. (**1984**): Does the Gettier Problem Rest on a Mistake? In *Mind, New Series* 93 (372), pp. 501–513.

Koch, M.;Richter, A.;Schlosser, A. (**2007**): Produkte zum IT-gestützten Social Networking in Unternehmen. In *Wirtschaftsinformatik* 49 (6), pp. 448–455.

Koenig, M. E. D. (**2004**): Knowledge Management Strategy: Codification Versus Personalization (A False Dichotomy). In Michael E. D. Koenig, T. Kanti Srikantaiah (Eds.):

Knowledge management lessons learned what works and what doesn't. 1st. Medford, N.J.: American Society for Information Science and Technology (ASIST monograph series), pp. 83–87.

Koenig, Michael E. D.; Srikantaiah, T. Kanti (Eds.) (**2004)**: Knowledge management lessons learned what works and what doesn't. 1st. Medford, N.J.: American Society for Information Science and Technology (ASIST monograph series). Available online at http://www.loc.gov/catdir/toc/ecip046/2003015556.html.

Kotonya, Gerald; Sommerville, Ian (**1998)**: Requirements engineering. Processes and techniques. Chichester ; New York: J. Wiley (Worldwide series in computer science).

KPMG (**2000)**: Knowledge management research report. KPMG.

Krcmar, H. (**2002)**: Wissensmanagement - nur eine Methode? In Matthias Bellmann, Helmut Krcmar, Tom Sommerlatte (Eds.): Praxishandbuch Wissensmanagement Strategien - Methoden - Fallbeispiele. 1st. Düsseldorf: Symposion Publ. (Wissensmanagement), pp. 26–28.

Krcmar, Helmut (**2010)**: Informationsmanagement. 5., vollst. überarb. und erw. Berlin, Heidelberg: Springer.

Kreibich, K. (**2003)**: Was machen oesterreichische Wissensmanagerinnen und Wissensmanager wirklich? Wissen & Management: Working Papers des Fachhochschul-Studiengangs Informationsberufe. Eisenstadt.

Kröll, M.;Rath, A.;Granitzer, M. et al.: Contextual Retrieval in Knowledge Intensive Business Environments. In : GI-Workshop Information Retrieval.

Krönig, Jan (**2001)**: Do incentive systems for knowledge management work? An empirical study on the design and influence of incentive systems of knowledge creation and transfer in manufacturing based industry. 1st. Bern: Lang (Europäische Hochschulschriften : Reihe 5, Volks- und Betriebswirtschaft ; 2803). Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=009475177&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Krulwich, B.;Burkey, C.;Consulting, A. (**1996)**: The ContactFinder agent: Answering bulletin board questions with referrals. Proceedings of the National Conference on Artificial Intelligence. In : Proceedings of the Thirteenth National Conference on Artificial Intelligence and Eighth Innovative Applications of Artificial Intelligence Conference. Proceedings of the Thirteenth National Conference on Artificial Intelligence : Eight Innovative Applications of Artificial Intelligence Conference (IAAI-96). Menlo Park, Calif: AAAI Press/The MIT Press.

Lakhani, K. R.;Hippel, E. von (**2003)**: How open source software works: “free” user-to-user assistance. In *Research Policy* 32 (6), pp. 923–943.

Laudon, Kenneth C.; Laudon, Jane Price (**2006)**: Management information systems managing the digital firm. 9th. Upper Saddle River, N.J.: Pearson Prentice Hall International. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=014180410&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Laudon, Kenneth C.; Laudon, Jane Price (**2009)**: Management information systems managing the digital firm. Upper Saddle River, NJ: Pearson Prentice Hall.

Laudon, Kenneth C.; Laudon, Jane Price (**2007)**: Management information systems managing the digital firm. 10th. Upper Saddle River, NJ: Pearson Prentice Hall.

Lave, J.; Wenger, E. (**1991)**: Situated learning: Legitimate peripheral participation: Cambridge university press.

Law, S.;Chang, M. (**2008)**: Fostering knowledge exchange in online communities: a social capital building approach. In : Proceedings of the International Conference of in Information Systems. ICIS. Paris, France, 14.-17.12.2008, pp. Paper 173.

Lehner, Franz (**2006)**: Wissensmanagement Grundlagen, Methoden und technische Unterstützung. 1st. München: Hanser. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=014573104&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=014573104&line_number=0002&func_code=DB_RECORDS&service_type=MEDIA Klappentext".

Lehner, Franz (**2008)**: Wissensmanagement Grundlagen, Methoden und technische Unterstützung. 2nd. München: Hanser. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=016332727&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Lehner, Franz (**2009)**: Wissensmanagement Grundlagen, Methoden und technische Unterstützung. 3rd. München: Hanser. Available online at "http://deposit.d-nb.de/cgi-bin/dokserv?id=3200362&prov=M&dok_var=1&dok_ext=htm Inhaltstext http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=017072733&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Leimeister, J. M. (**2004)**: Pilotierung Virtueller Gemeinschaften im Gesundheitswesen: Pilotierung virtueller Communities im Gesundheitsbereich - Bedarfsgerechte Entwicklung, Einführung und Betrieb.

Liebowitz, J. (**1999)**: Key ingredients to the success of an organization's knowledge management strategy. In *Knowledge and Process Management* 6 (1), pp. 37–40. Available online at "http://dx.doi.org/10.1002/(SICI)1099-1441(199903)6:1<37::AID-KPM40>3.0.CO;2-M".

Lin, N. (**2001)**: Building a network theory of social capital. In *Social capital: Theory and research*, pp. 3–29.

Linden, G.;Smith, B.;York, J. (**2003)**: Amazon. com recommendations: Item-to-item collaborative filtering. In *IEEE Internet computing* 7 (1), pp. 76–80.

Lokaiczyk, R.;Godehardt, E.;Faatz, A. et al. (**2007)**: Exploiting context information for identification of relevant experts in collaborative workplace-embedded e-learning environments. In Erik Duval, Ralf Klamma, Martin Wolpers (Eds.): Creating new learning experiences on a global scale. Second European Conference on Technoloy Enhanced Learning, EC-TEL 2007, Crete, Greece, September 2007 ; proceedings. European Conference on Technology Enhanced Learning; EC-TEL; European Conference on Technoloy Enhanced Learning. Berlin: Springer (Lecture notes in computer science, 4753), pp. 217–231.

Lullies, V.;Bollinger, H.;Weltz, F. (**1993)**: Wissenslogistik. Gesamthochsch, Frankfurt am Main.

Lundvall, B.;Johnson, B. (**1994)**: The learning economy. In *Industry & Innovation* 1 (2), pp. 23–42.

Ma, D.;Schuler, D.;Zimmermann et al. (**2009)**: Expert Recommendation with Usage Expertise. In : Proceedings of the 25th IEEE International Conference on Software Maintenance. ICSM 2009. Edmonton, Alberta, Canada, September 20-26, 2009: IEEE.

Maanen, J. van;Barley, S. (**1984)**: Occupational Communities: Culture and control in organizations. In *Research in Organizational Behavior* (6), pp. 187–365.

Maier, Ronald (**2007)**: Knowledge management systems information and communication technologies for knowledge management ; with 91 tables. 3rd. Berlin: Springer. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=015632030&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=015632030&line_number=0002&func_code=DB_RECORDS&service_type=MEDIA Klappentext".

Makolm, J.;Weiß, S.;Reisinger, D. (**2007**): Proactive knowledge management: the DYONIPOS research and use-case project. In : Proceedings of the 1st international conference on theory and practice of elecontric governance. ICEGOV. New York, NY, USA: ACM.

March, S.;Smith, G. (**1995**): Design and natural science research on information technology. In *Decision Support Systems* 15 (4), pp. 251–266.

Markus, M. L. (**2001**): Toward a Theory of Knowledge Reuse: Types of Knowledge Reuse Situations and Factors in Reuse Success. In *Journal of Management Information Systems* 18, pp. 57–93. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=4753192&site=ehost-live".

Mårtensson;Maria (**2000**): A critical review of knowledge management as a management tool. In *Journal of Knowledge Management* 4 (3), pp. 204–216.

Maybury, M. T. (**2006**): Expert finding systems. In *Bedford, Massachusetts, The MITRE Corporation*.

McAdam, R.;McCreedy, S. (**1999**): A critical review of knowledge management models. In *The Learning Organization* 6 (3).

McCarthy, John (**1963**): Situations, actions, and causal laws: Stanford University, Artificial Intelligence Project.

McCarthy, J. (**2002**): Actions and Other Events in Situation Calculus. In Dieter Fensel, Fausto Giunchiglia and Deborah L. McGuinness and Mary-Anne Williams (Eds.): Proceedings of the Eights International Conference on Principles and Knowledge Representation and Reasoning (KR-02). KR-02. Toulouse, France, April 22-25, 2002: Morgan Kaufmann.

Medeiros, A. K. A. de;Dongen, B. F. van;Aalst, W. van der et al. (**2004**): Process Mining: Extending the \alpha-algorithm to Mine Short Loops. Eindhoven University of Technology, Eindhoven.

Medeiros, A.K Alves de; Weijters, A.J.M.M; Aalst, W.M.P. van der (**2005a**): Using genetic algorithms to mine process models. Representation, operators and results. Eindhoven: [Beta, Research School for Operations Management and Logistics]. Available online at http://www.worldcat.org/oclc/66640289.

Medeiros, A. de;Dongen, B. van;Aalst, W. van der et al. (**2005b**): Process mining for ubiquitous mobile systems: an overview and a concrete algorithm. In *Ubiquitous Mobile Information and Collaboration Systems*, pp. 151–165.

Medeiros, A. de;Weijters, A.;Aalst, W. van der (**2005c**): Genetic Process Mining: A Basic Approach and Its Challenges. In Christoph Bussler, Armin Haller (Eds.): Business Process Management Workshops. Business Process Management Workshops. Berlin: Springer (Lecture notes in computer science, 3812), pp. 203–215.

Mertins, Kai (**2003**): Knowledge management concepts and best practices. 2nd. Berlin: Springer.

Miller, B. N.;Albert, I.;Lam, S. K. et al.: MovieLens unplugged: experiences with an occasionally connected recommender system. Proceedings of the 8th international conference on Intelligent user interfaces.

Moffett, S.;McAdam, R.;Parkinson, S. (**2002**): Developing a model for technology and cultural factors in knowledge management: a factor analysis. In *Knowledge and Process Management* 9 (4), pp. 237–255. Available online at "http://dx.doi.org/10.1002/kpm.152".

Nahapiet, J.;Ghoshal, S. (**1998**): Social capital, intellectual capital, and the organizational advantage. In *Academy of management review* 23 (2), pp. 242–266.

Naisbitt, John (**1982**): Megatrends. Ten new directions transforming our lives. New York: Warner Books.

Nerdinger, Friedemann W. (**1995**): Motivation und Handeln in Organisationen. Eine Einführung. Stuttgart: W. Kohlhammer.
Nerdinger, Friedemann W. (**2003**): Motivation von Mitarbeitern. Göttingen: Hogrefe.
Nerdinger, F. W. (**2004**): Die Bedeutung der Motivation beim Umgang mit Wissen. In Gabi Reinmann, Heinz Mandl (Eds.): Psychologie des Wissensmanagements. Perspektiven, Theorien und Methoden. 1.th ed. Göttingen: Hogrefe, pp. 91–101.
Newell, A.; Simon, H.A (**1972**): Human problem solving: Prentice-Hall Englewood Cliffs, NJ (104).
Ngu, Anne H. H.; Kitsuregawa, Masaru; Neuhold, Erich J.; Chung, Jen-Yao; Sheng, Quan Z.; Dubouloz, Benoit; Toklu, Candemir (Eds.) (**2005**): Discovering the Most Frequent Patterns of Executions in Business Processes Described in BPEL: Springer (Lecture notes in computer science, 3806).
Nonaka, I. (**1994**): A dynamic theory of organizational knowledge creation. In *Organization science* 5 (1), pp. 14–37.
Nonaka, Ikujir; Toyama, Ryoko; Hirata, Tour (**2008**): Managing flow a process theory of the knowledge-based firm. 1st. New York: Palgrave Macmillan. Available online at http://www.loc.gov/catdir/enhancements/fy0829/2008016169-t.html.
Nonaka, Ikujiro (**2001**): Knowledge emergence social, technical, and evolutionary dimensions of knowledge creation. Oxford: Oxford Univ. Press. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=009380833&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".
Nonaka, Ikujiro; Takeuchi, Hirotaka (**1997**): Die Organisation des Wissens wie japanische Unternehmen eine brachliegende Ressource nutzbar machen. 1st. Frankfurt: Campus Verl. (The knowledge creating company «dt.»). Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=007458575&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".
Nonaka, Ikujirō; Takeuchi, Hirotaka (**1995**): The knowledge creating company. How Japanese companies create the dynamics of innovation. New York: Oxford Univ. Press. Available online at http://www.gbv.de/dms/faz-rez/F19980518BONUS--100.pdf.
North, K.;Romhardt, K.;Probst, G. (**2000**): Wissensgemeinschaften. Keimzellen lebendigen Wissensmanagements. In *IO Management* 69 (7), pp. 52–62.
North, Klaus (**2005**): Wissensorientierte Unternehmensführung Wertschöpfung durch Wissen. 4th. Wiesbaden: Gabler. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=014281343&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Klappentext".
North, K.;Varlese, N. (**2006**): Motivieren für die Wissensteilung und Entwicklung (26.01). Available online at http://www.wissensmanagement.net/online/archiv/2001/02_0301/wissensmanagement_anreize.shtml.
Nunamaker, J.;Chen, M.;Purdin, T. (**1991**): Systems Development in Information Systems Research. In *Journal of Management Information Systems* 7 (3), pp. 89–106.
O'Dell, C.;Grayson, C. J. (**1998**): If Only We Knew What We Know: Identification and Transfer of Internal Best Practices. In *California Management Review* 40, pp. 154–174. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=738862&site=ehost-live".
Oldigs-Kerber, J.;Shpilsky, A.;Sorensen, S. (**2003**): The Aventis approach to knowledge management: Locating inhouse expertise for immediate anytime, anywhere availability. In K. Mertins, P. Heisig, J. Vorbeck (Eds.): Knowledge management: concepts and best practices: Springer Verlag, pp. 262–272.

Oldigs-Kerber, J. (2007): Aspekte zum Wissensmanagement: Wissen-teilen als soziales Dilemma. In *Wissensmanagement* 49 (1), pp. 65–70.

Olivera, F.;Goodman, P.;Tan, S. (2008): Contribution behaviors in distributed environments. In *Mis Quarterly* 32 (1), pp. 23–42.

Orlikowski, W.;Hoffman, D. (1997): An Imporvisational Model for Change Managment: The Case of Groupware Technologies. In *Sloan Management Review* 38 (2), pp. 11–21.

Österle, H.;Becker, J.;Frank, U. et al. (2010): Memorandum zur gestaltungsorientierten Wirtschaftsinformatik. In Hubert Österle, Robert Winter, Walter Brenner (Eds.): Gestaltungsorientierte Wirtschaftsinformatik. Ein Plädoyer für Rigor und Relevanz. Nürnberg: Infowerk, pp. 1–6.

Peterson, James Lyle (1981): Petri net theory and the modeling of systems. Englewood Cliffs, N.J: Prentice-Hall.

Petter, C.;Reich, K.;Helling, K. (2007): Social Software and the Establishment of Virtual Communities of Practice in the Tourism Sector. In *eLearning Papers* (5). Available online at http://www.elearningpapers.eu/index.php?page=doc&doc_id=10223&doclng=6.

Pfeffers, K.;Tuunanen, T.;Gengler, C. et al. (2006): The Design science research process: A model for producing and presenting information systems research. In : Proceedings of the 1st International Conference on Design Science Research in Information Systems and Technology (DESRIST 2006). DESRIST. Claremont, CA, USA, 24-25 Feb 2006: CGU, pp. 83–106.

Piaget, Jean (1996): Logique et connaissance scientifique. Paris: Gallimard (Encyclopédie de la Pléiade, 22).

Pinsonneault, A.;Heppel, N. (1997): Anonymity in group support systems research: a new conceptualization, measure, and contingency framework. In *Journal of Management Information Systems* 14 (3), pp. 89–108.

Platon (n.d.): Theaitetos.

Polanyi, Michael (1967): The tacit dimension. Garden City, N.Y: Doubleday. Available online at http://www.worldcat.org/oclc/222979784.

Probst, Gilbert; Raub, Steffen; Romhardt, Kai (2000): Managing knowledge building blocks for success. 1st. Chichester: Wiley (Wissen managen «engl.»). Available online at http://www.loc.gov/catdir/toc/onix02/99037215.html.

Probst, Gilbert; Raub, Steffen; Romhardt, Kai (2006): Wissen managen wie Unternehmen ihre wertvollste Ressource optimal nutzen. 5th. Wiesbaden: Gabler. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=014756193&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Klappentext".

Prusak, L. (2001). In *IBM Systems Journal* 40 (4), pp. 1002–1007.

Quinn, James Brian (1992): Intelligent enterprise. A knowledge and service based paradigm for industry. New York: Free Press. Available online at http://www.loc.gov/catdir/description/simon032/92019384.html.

Qureshi, S.;Briggs, R.;Hlupic, V. (2006): Value Creation from Intellectual Capital: Convergence of Knowledge Management and Collaboration in the Intellectual Bandwidth Model. In *Group Decision and Negotiation* 15 (3), pp. 197–220. Available online at http://dx.doi.org/10.1007/s10726-006-9018-x.

Rambow, R.;Bromme, R. (2000): Was Schöns" reflective practitioner" durch die Kommunikation mit Laien lernen könnte. In *Wissen-Können-Reflexion: Ausgewählte Verhältnisbestimmungen*, pp. 245–263.

Rath, A.;Kröll, M.;Andrews, K. et al. (**2006**): Synergizing standard and ad-hoc processes. In *Practical Aspects of Knowledge Management*, pp. 267–278.

Rehäuser, J.;Krcmar, H. (**1996**): Wissensmanagement im Unternehmen. In *Managementforschung* 6.

Reich, Jochen (**2008**): Supporting the execution of knowledge intensive processes by means of expert and best-practice mediation. 1st ed. München: Dr. Hut.

Reiß, M. (**1997**): Change Management als Herausforderung. In Michael Reiß, Lutz von Rosenstiel, Annette Lanz (Eds.): Change Management: Programme, Projekte und Prozesse. Stuttgart, pp. 5–29.

Reiter, R. (**2001**): Knowledge in action: logical foundations for specifying and implementing dynamical systems: The MIT Press.

Resnick, P.;Varian, H. R. (**1997**): Recommender systems. In *Communications of the ACM* 40 (3), p. 58.

Richter, A.;Koch, M. (**2008**): Wissenssuche: Wo sind die Experten? In *Wissensmanagement* (5), pp. 18–20.

Riempp, Gerold (**2004**): Integrierte Wissensmanagement-Systeme Architektur und praktische Anwendung. 1st. Berlin: Springer (Business-Engineering). Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=010601977&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

Roberts, Royston M. (**1989**): Serendipity. Accidental discoveries in science. New York: Wiley (Wiley science editions). Available online at http://www.loc.gov/catdir/description/wiley034/88033638.html.

Robertson, M.;O'Malley Hammersley, G. (**2000**): Knowledge management practices within a knowledge-intensive firm: the significance of the people management dimension. In *Journal of European Industrial Training* 24 (2/3/4), pp. 241–253.

Rockart, J. F. (**1979**): Chief executives define their own data needs. In *Harvard Business Review* 57, pp. 81–93. Available online at "http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=3867668&site=ehost-live".

Rollett, Herwig (**2003**): Knowledge management processes and technologies. 1st. Boston: Kluwer Acad. Publ.

Rosenstiel, Lutz von (**2007**): Grundlagen der Organisationspsychologie. Basiswissen und Anwendungshinweise. 6., überarb. Stuttgart: Schäffer-Poeschel.

Schmidl, J.;Slavtchev, V.;Wittges, H. et al. (**2011a**): Knowledge Management Success or Failure - What Determines the Performance of a KM-Initative? In Ronald Maier (Ed.): Proceedings der 6. Konferenz Professionelles Wissensmanagement. Vom Wissen zum Handeln. Professionelles Wissensmanagement. Innsbruch, Österreich, February, 21.-23. 2011.

Schmidl, J.;Wittges, H.;Krcmar, H. (**2011b**): Similarity Determination in Activity Secquences - A Supportive Framework. In Abraham Bernstein, Gerhard Schwabe (Eds.): Proceedings of the 10th International Conference on Wirtschaftsinformatik - Volume 2 of 2. Wirtschaftsinformatik. Zürich, Schweiz, February, 16-18 2011, pp. 797–807.

Schmidt, K. (**2000**): The critical role of workplace studies in. In Paul Luff, Jon Hindmarsh, Christian Heath (Eds.): Workplace studies. Recovering work practice and informing system design. Cambridge: Cambridge university press, pp. 141–153.

Schneider, Klaus; Schmalt, Heinz-Dieter (**2000**): Motivation. With assistance of Klaus Schneider. 3., überarb. und erw. Stuttgart: Kohlhammer (Basisbuch).

Schuler, D.;Zimmermann, T. (**2008**): Mining usage expertise from version archives. In : Proceedings of the International Conference on Software Engineering & co-located work-

shops. Leipzig, Germany,, May 10 - 18, 2008. Working Conference on Mining Software Repositories; MSR. Available online at http://doi.acm.org/10.1145/1370750.1370779.
Schütt, P. (**2003**): The post-Nonaka knowledge management. In *Journal of Universal Computer Science* 9 (6), pp. 451–462.
Schwarz, S.;Roth-Berghofer, T. (**2003**): Towards goal elicitation by user observation. In Andreas Hotho, Gerd Stumme (Eds.): Proceedings of the LLWA 2003. LLWA, 10/2003. Karlsruhe: GI, pp. 224–228.
Schwarzer, Bettina; Krcmar, Helmut (**2004**): Wirtschaftsinformatik. Grundzüge der betrieblichen Datenverarbeitung. 3., überarb. Stuttgart: Schäffer-Poeschel.
Seid, D. Y.;Kobsa, A. (**2003**): Expert finding systems for organizations: Problem and domain analysis and the DEMOIR approach. In *JOURNAL OF ORGANIZATIONAL COMPUTING AND ELECTRONIC COMMERCE*.
Serdyukov, P. (**2009**): Search for expertise. Going beyond direct evidence (SIKS dissertation series).
Serdyukov, P.;Feng, L.;van Bunningen, A. et al. (**2008**): The Right Expert at the Right Time and Place. In *Practical Aspects of Knowledge Management*, pp. 38–49.
Shkundina, R.;Schwarz, S. (**2005**): A similarity measure for task contexts. In : Proceedings of the Workshop Similarities-Processes-Workflows in conjunction with the 6th International Conference on Case-Based Reasoning.
Silver, M.;Markus, M.;Beath, C. (**1995**): The information technology interaction model: a foundation for the MBA core course. In *Mis Quarterly*, pp. 361–390.
Simon, Herbert Alexander (**1996**): The sciences of the artificial. 3rd ed. Cambridge Mass. u.a: MIT Pr.
Simon, H. A. (**2002**): Organizing and coordinating talk and silence in organizations. In *Industrial and Corporate Change* 11 (3), pp. 611–618.
Skyrme, D.;Amidon, D. (**1997**): The Knowledge Agenda. In *Journal of Knowledge Management* 1 (1), pp. 27–37.
Slavtchev, V. (2009): Success Factors in Knowledge Management Projects – a Literature Analysis.
Soliman, F.;Spooner, K. (**2000**): Strategies for implementing knowledge management: role of human resources management. In *Journal of Knowledge Management* 4 (4), pp. 337–345.
Stahlknecht, Peter; Hasenkamp, Ulrich (**2005**): Einfuhrung in die Wirtschaftsinformatik. 11.th ed. Berlin ; New York: Springer. Available online at http://www.myilibrary.com?id=61589.
Staw, B. (**1981**): The escalation of commitment to a course of action. In *Academy of management review* 6 (4), pp. 577–587.
Stonehouse, G. H.;Pemberton, J. D. (**1999**): Learning and knowledge management in the intelligent organisation. In *Participation and Empowerment: An International Journal* 7 (5), pp. 131–144.
Sukowski, Oliver (**2002**): Knowledge management case studies project experiences, implementation insights, key questions. 1st. St. Gallen: NetAcademy Press. Available online at "http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=009621544&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".
Tajfel, H. (**1986**): JC (1986). The social identity theory of intergroup behavior. In *Psychology of intergroup relations*, pp. 7–24.
Tedmori, S.;Jackson, T. W.;Bouchlaghem, D. (**2006**): Locating knowledge sources through keyphrase extraction. In *Journal of Knowledge and Process Management* 13 (2), p. 7.

Teigland, R. (**2003**): Knowledge networking: Structure and performance in networks of practice: Institute of International Business, Stockholm School of Economics [Institutet för internationellt företagande vid Handelshögsk.](IIB).

Teigland, R.;Wasko, M. (**2006**): What are Electronic Networks of Practice and How Do These Online Social Structures Differ from Communities of Practice? In S. Dasgupta (Ed.): Encyclopedia of virtual communities and technologies: IGI Global.

Terveen, L.;Hill, W. (**2002**): Beyond Recommender Systems: Helping People Help Each Other. In John Millar Carroll (Ed.): Human-computer interaction in the new millennium. New York: ACM Press.

Thiedeke, U. (**2003**): Virtuelle Gruppen: Westdeutscher Verlag GmbH.

Thompson, C. A.;Göker, M. H.;Langley, P. (**2004**): A personalized system for conversational recommendations. In *Journal of Artificial Intelligence Research* 21 (1), pp. 393–428.

Torraco, R. J. (**2005**): Writing Integrative Literature Reviews: Guidelines and Examples. In *Human Resource Development Review* 4 (3), pp. 356–367. Available online at "http://hrd.sagepub.com/cgi/content/abstract/4/3/356" "10.1177/1534484305278283".

Tsoukas, H. (**1998**): Forms of knowledge and forms of life in organized contexts. In *the Realm of Organization: Essays for Robert Cooper*, pp. 43–66.

Turner, J.C; Hogg, M.A; Oakes, P.J; Reicher, S.D; Wetherell, M.S (**1987**): Rediscovering the social group: A self-categorization theory: Basil Blackwell.

Vanderfeesten, I. T. P.;Reijers, H. A.;Aalst, W. M. P. van der (**2008**): Product Based Workflow Support: Dynamic Workflow Execution. In Zohra Bellahsène, Michel Léonard (Eds.): Advanced information systems engineering. 20th international conference, CAiSE 2008, Montpellier, France, June 16-20, 2008 ; proceedings. Berlin: Springer (Lecture notes in computer science, 5074), pp. 571–574.

Venkatesh, V.;Morris, M. G.;Davis, G. B. et al. (**2003**): User acceptance of information technology: Toward a unified view. In *Management Information Systems Quarterly* 27 (3), pp. 425–478.

Verbeek, H. M.;Buijs, J.;van Dongen, B. F. et al. (**2010**): ProM 6: The Process Mining Toolkit. In Marcello La Rosa (Ed.): Proceedings of the Business Process Management 2010 Demonstration Track. BPMDemos 2010. Hoboken, USA,, September 14-16, 2010.

Vitari, C.;Ravarini, A. (**2009**): The Effects Of Communities Of Practice On The Success Of An Expert Recommending Service. In *Grenoble Ecole de Management (Post-Print).*

Völker, Rainer; Sauer, Sigrid; Simon, Monika (**2007**): Wissensmanagement im Innovationsprozess mit … 17 Tabellen. 1st. Heidelberg: Physica-Verl. Available online at "http://deposit.ddb.de/cgi-bin/dokserv?id=2753795&prov=M&dok_var=1&dok_ext=htm Inhaltstext http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&doc_number=015025020&line_number=0001&func_code=DB_RECORDS&service_type=MEDIA Inhaltsverzeichnis".

W3C (**1999**): XML Path Language (XPath) Version 1.0. Available online at http://www.w3.org/TR/xpath/, checked on 21/01/2011.

Wasko, M.;Faraj, S. (**2000**): "It is what one does": why people participate and help others in electronic communities of practice. In *Journal of strategic information systems* 9 (2-3), pp. 155–173.

Wasko, M.;Faraj, S. (**2005**): Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. In *Mis Quarterly* 29 (1), pp. 35–57.

Webster, J.;Watson, R. T. (**2002**): Analyzing the past to prepare for the future: Writing a literature review. In *MIS Quarterly* 26 (2), pp. xiii–xxiii.

Wegner, D. M. (**1987**): Transactive memory: A contemporary analysis of the group mind. In Brian Mullen, G. R. Goethals (Eds.): Theories of group behavior. New York, Berlin, Heidelberg, London, Paris, Tokyo: Springer (Springer series in social psychology), pp. 185–205.
Wegner, D. M.;Giuliano, T.;Hertel, P. (**1985**): Cognitive interdependence in close relationships. In William John Ickes (Ed.): Compatible and incompatible relationships. New York, NY: Springer (Springer series in social psychology), pp. 253–276.
Wen, L.;Wang, J.;Aalst, W. van der et al. (**2004**): A Novel Approach for Process Mining Based on Event Types. Eindhoven University of Technology, Eindhoven.
Wenger, E. (**1998**): Communities of practice: Learning as a social system. In *Systems thinker* 9 (5), pp. 1–5.
Wenger, E. C.;Snyder, W. M. (**2000**): Communities of practice: The organizational frontier. In *Harvard business review* 78 (1), pp. 139–146.
Wenger, Etienne; McDermott, Richard A.; Snyder, William (**2002**): Cultivating communities of practice. A guide to managing knowledge. Boston, Mass: Harvard Business School Press. Available online at http://edrev.asu.edu/reviews/rev405.htm.
WFMC: Introduction to the Workflow Management Coalition. Available online at http://www.wfmc.org/about.htm, checked on 21/01/2011.
WFMC (**1998**): Audit data specification. Technical Report WFMCTC - 1015 Version 1.1.
Wiig, K. M. (**1997**): Knowledge Management: Where Did It Come From And Where Will It Go? In *Expert Systems With Applications* 13 (1), pp. 1–14.
Willke, H. (**1998**): Organisierte Wissensarbeit. In *Zeitschrift für Soziologie* 27 (3), pp. 161–177.
Wilson, T. (**2002**): The nonsense of knowledge management. In *Information research* 8 (1), pp. 8-1.
Witten, I. H.;Frank, E. (**2002**): Data mining: practical machine learning tools and techniques with Java implementations. In *ACM SIGMOD Record* 31 (1), pp. 76–77.
Wong, K. Y. (**2005**): Critical success factors for implementing knowledge management in small and medium enterprises. In *Industrial Management & Data Systems* 105 (3), pp. 261–279. Available online at "10.1108/02635570510590101".
Wong, K. Y.;Aspinwall, E. (**2005**): An empirical study of the important factors for knowledge-management adoption in the SME sector. In *Journal of Knowledge Management* 9 (3), pp. 64–82.
Wong, K. Y.;Aspinwall, E. (**2004**): Characterizing knowledge management in the small business environment. In *Journal of Knowledge Management* 8 (3), pp. 44–61.
Wu, L.;Waber, B.;Aral, S. et al. (**2008**): Mining face-to-face interaction networks using sociometric badges: Predicting productivity in an it configuration task. In : Proceedings of the International Conference of in Information Systems. ICIS. Paris, France, 14.-17.12.2008, pp. Paper 127.
Yahya, S.;Goh, W.-K. (**2002**): Managing human resources toward achieving knowledge management. In *Journal of Knowledge Management* 6 (5), pp. 457–468.
Ye, Y.;Nakakoji, K.;Yamamoto, Y. (**2008**): Understanding and Improving Collective Attention Economy for Expertise Sharing. In Zohra Bellahsène, Michel Léonard (Eds.): Advanced information systems engineering. 20th international conference, CAiSE 2008, Montpellier, France, June 16-20, 2008 ; proceedings. Berlin: Springer (Lecture notes in computer science, 5074), pp. 167–181.
Yetim, F. (**2009**): From Communicative Action Theory to Socio-Technical Artifacts: Presentation of Three System Prototypes. In Jörg Becker, Helmut Krcmar, Björn Niehaves (Eds.):

Wissenschaftstheorie und gestaltungsorientierte Wirtschaftsinformatik. Heidelberg: Physica-Verl., pp. 23–42.

Yu-Chung Hung;Shi-Ming Huang;Quo-Pin Lin et al. (**2005)**: Critical factors in adopting a knowledge management system for the pharmaceutical industry. In *Industrial Management & Data Systems* 105 (2), pp. 164–183. Available online at "10.1108/02635570510583307".

Zack, M. H. (**1999)**: Developing a Knowledge Strategy. In *California Management Review* 41, pp. 125–145. Available online at "http://search.ebscohost.com/login.aspx?direct=true &db=buh&AN=1867074&site=ehost-live".

Zeleny, Milan (**2005)**: Human systems management integrating knowledge management and systems. 1st. New Jersey: World Scientific (HSM).

Ziegler, C.-N.;Golbeck, J. (**2007)**: Investigating Interactions of Trust and Interest Similarity. In *Decision Support Systems* 43 (2), pp. 460-475.

Zimbardo, P. G. (**1983)**: Psychologie. Anhang: Lernhilfen und Arbeitshilfen von K. Westhoff ; mit dreihundertzweiundzwanzig zum Teil farbigen Abbildungen. 4. neubearbeitete Auflage. Berlin: Springer.

Appendix A XML Application for Data Transformation

Appendix A-1 Transformations in General

A transformation configuration consists of a list of different map entries. Each *mapEntry* poses a certain condition on the data and defines a target of the transformation, that is inserted in the output, if the condition is met by the data. There are two kinds of transformation definitions. The first one details how to map data changes to activities directly, while the second uses other transformation definitions, that can in turn depend on other map entries. By distinguishing between these two kinds, and by allowing to transitively use other map entries, it is obvious that a hierarchy of transformation definitions can be build. Each single transformation entry is associated to a hierarchy level by using the attribute *hierachyLevel*. A transformation that details how to map, without relying on other transformation definitions, is on hierarchy level one. If a transformation utilizes information from other transformations it is on a hierarchy level that is the level of the highest used transformation plus one. If e.g. a transformation entry uses a map entry of hierarchy level one and another that has hierarchy level three, its own level is four. In the following the two different ways of transforming data, the basic transformation and the dependent transformation are described.

Appendix A-2 Basic Transformations

The basic transformation, which defines how to map data in a straight forward manner, has a format as shown in the following. Essentially a basic transformation consists of two obligatory elements, the *activity* node and the *gets* node as well as one optional element, the *attribute* node.

```
<mapEntry hierachyLevel="1">
  <activity name="the_identified_activity"
     status="the_associated_state"/>
     <gets>
        <get name="name_of_getter" type="xPath"
        xPath="self::node()[parent::History]/..." />
     </gets>
     <attributes>
        <get type="xPath" xPath="self::node()/@oldValue"
        name="oldValue" />
        <get type="xPath" xPath="self::node()/@newValue"
        name="newValue" />
        <derived>
          <derived name="change" type="conditonalResult">
            <condition type="less">
              <get ref="oldValue" type="refGet">
                <operator type="toInteger" />
```

```
                </get>
                <get ref="newValue" type="refGet">
                  <operator type="toInteger" />
                </get>
              </condition>
              <true result="increased" />
              <false result="deceased" />
            </derived>
          </derived>
       </attributes>
    </activity>
  </mapEntry>
```

Appendix A-2-1 The gets Node

The gets node defines how to select a value from the input document and by this means determines the pattern that is checked for. To perform this task, a gets node consists of nested get nodes. The get nodes offer a highly customizable way of accessing data that e.g. resides in a database or an XML file. In all cases of the transformation, the data retrieval was done by applying an XPath Query against the document, as indicated by the type attribute, due to the fact that only XML-coded data is accessed. However, querying a database instead of XML coded data would change only the get node's content while the rest of the configuration could stay unchanged. XPath queries are used for two complementary reasons. First, the XPath is used to check if the data fits a certain pattern. If the XPath expression does yield an empty result, when applied to the document, thus the provided pattern does not exist in the document, the map entry's transformation will not be applied. The XPath expression is therefore used as a selector. On the other hand, if the XPath did select a value, thus the desired pattern is present in the source document, the map entry's transformation is applied and therefore the output document will contain this entry's activity. In the second case, e.g. whenever the pattern is found, the XPath serves another purpose by providing access to the data at the specified location. This fact is exploited in the third, optional part of a basic map entry to extract attribute information as described shortly.

Appendix A-2-2 The activity Node

The second mandatory node, the activity node, determines the activity that is associated with this map entry, in case the application of the gets node did return a non-empty value. The transformation output will therefore contain an activity entry with the name of the attribute name and a state transition that corresponds to the status attribute. The possible state transitions are those of the MXML format as shown in listing 4.2, while the activity name can be chosen arbitrarily, but should convey the semantic meaning of the data state change.

Appendix A-2-3 The attributes Node

The optional attributes node, can be used to extract additional information from the source document. The extracted information will be stored into the MXML format's data node of the respective audit trail entry. The extraction can be one of two types. Either by using a get node, it is possible to query the document to determine an attribute value directly. As data is XML-coded data, a get node of type XPath will be used again. In this case, the XPath expression is evaluated against the XPath query result of the get node one level higher. Therefore the get node that was used to check if the input document did comply to the expected pattern defines the context of the attributes node's get nodes. The name attribute of the get node, defines the new data elements name in the MXML data node. The second way of extracting an attribute is by dynamically determining it. This is achieved by using the information of the source document to construct a derived attribute. A derived node is used for this purpose, as also shown in the example fragment in the listing above. The derived node contains one node for each attribute that is to be dynamically generated, wrapped into a node that is also named derived. The attribute name of the later one, determines the name of the node, which will be inserted into the MXML format's data node of the respective audit trail entry. The type attribute determines how to create the value. In the example fragment, a conditional assignment for creating the new attribute is used. To do so, a condition is posed, that is checked on the data by using the condition node. There is a wide variety of conditions that can be used. The example uses a "less" condition that checks if the first value is less than the second value. consistent with the rest of the highly customizable configuration, the values to be compared are found by using get nodes, as described above already, and operators to transform the result in a suitable format. There is a great variety of conditions that compare two elements, or check on one element alone. The depicted derived node is of type "conditionalresult" and assigns either the value of the true node's result attribute or the value of the false node's result attribute to the newly created attribute, depending on the result of the conditional check. In the example fragment, the newly created attribute "change" therefore has either value "increased" or value "decreased", if the new value is higher than the old one for the first case and the other way around for the second case. Note that there are other ways of creating new and therefore derived attributes.

Appendix A-3 Hierarchical Transformations

The dependent transformation, which reuses already defined transformations, has a format that is shown in the listing below. Essentially a dependent transformation consists of two obligatory elements, the activity node and the depends node as well as one optional element, the attribute node. The activity node and the attribute node have the same syntax and semantics as they have in the basic transformation definitions explained in Appendix A-2. The second obligatory element is different however and will be described in the following.

Appendix A-3-1 The depends Node

The depends node is used to specify a transformation that is based upon other already defined transformations. It consists of one or more sonNode nodes and one condition node. The later one itself consists of one or more condition nodes that have the same syntax and semantics as the conditions used in e.g. a conditional result as explained in the section before. The get node used for determining the condition nodes input, is a get node of type “xpath. However it uses a new construct, the context attribute. This attribute specifies, that the query context and thus the start point for evaluating the XPath query contained in the XPath attribute is determined by the result of the referenced get node’s result. This way it is possible to reuse the results of transformations of lower hierarchy levels.

Each sonNode node represents one dependency on another transformation that is identified by the attribute ref. Note that the referenced name consists of the referenced transformation’s activity attribute content, an underscore and the referenced transformation’s status attribute content. This is necessary since one activity can have more than one state transition, although only one transition, namely “complete” will be used in the case study.

Whenever a transformation depends on another transformation, there is the question whether or not the dependent transformation should replace the transformation it depends on. The answer to this question depends on the way one wants to use the configuration. Basically there are three possible alternatives that are communicated by using the attribute replaceChild, can have the values “always”, “never” or “ifmatches”. To understand this distinction it is necessary to know how transformations are created and evaluated within the transformation application. The application searches the configuration file for map entries and fills its internal lists of configurations accordingly. It does so by starting with the lowest possible hierarchy level, working its way up the hierarchical definitions. When processing however, the application first evaluates the highest hierarchy levels first, eventually working its way down. Knowing this, the first possible attribute value “always”, specifies that the transformation upon which the current transformation depends, will never be used independently but only when referenced, by internally setting a certain flag in the lower level map entry. It is always replaced, thus the referenced node’s activity will never be contained in the output. This can be necessary and helpful for modularization. In case many map entries share a common pattern that the input should comply to, it makes sense to extract this common pattern into a mapEntry that cannot be evaluated on its own but can be “included” into the transformation that need it.

If the attribute value is “never”, the child node, will always also be evaluated and will eventually be contained in the output, if its matches. Thus in case the transformation that depends on other nodes does match, its child nodes will certainly also match, as their condition must have been fulfilled for the dependent transformation already. They will therefore also be integrated into the result. This option is necessary if the more specific result, represented by the dependent transformation entry as well as the transformation entries it depends on, need to be included into the result. An example might be that one wants to record that a process changed its status, but if it changed its status to a certain state, say its final state, then additionally another activity is identified, like e.g. “rapid progress” in case the process advanced many steps at once. Hence there are two activities that are inserted into the output. The last option for the child replacement attribute, the “ifmatches” value, can be used if the dependent transformation and its child node both convey a semantically valid piece of information. However the dependent transformation conveys more information and includes all the necessary information of the one it depends on. Therefore only the more information-rich dependent transformation should be included, if it is applicable and the less information rich transformation, in case the dependent transformation cannot be applied. An example situation for using this kind of child replacement is e.g. the need to determine that a process state changed, but if certain additional conditions hold, this event should not be considered as a process state change anymore, but rather as e.g. a process completion. Hence one of the two map entries activities is in the output. Note that both, the “never” and the “ifmatches” values can be used to express transformation that specialize or aggregate other transformations. The “never” option will increase information by adding additional information into the result and the “ifmatches” option will increase the information content by semantically enriching one element only.

Appendix B Documents Used and Created in First Case Study

Appendix B-1 Instruction Notes Delivered to Target Group

Experten finden – Expertise teilen: Wie komme ich besser voran in meinem Projekt

Um was geht’s?

Die Arbeit in neuen Umgebungen mit neuen Funktionen ist nie leicht. Bei SAP Systemen trifft dies ebenso zu, wie Sie sicher selbst nachvollziehen konnten. In besonderem Maße aber ist dies der Fall, wenn man das System nicht nur als Endnutzer zu Gesicht bekommt, sondern wenn man das System programmiert, konfiguriert oder allgemein erweitert – so wie Sie es tun. Hier sind nicht nur die Handhabung des Systems zu erlernen, sondern auch die Konzepte und Verknüpfungen im Hintergrund zu verstehen. Ihr Betreuer hilft Ihnen hier wo es geht. Eine weitere Möglichkeit Hilfe zu bekommen liegt im aktiven Austausch mit Ihren Kommilitonen. Während dies im Team naturgemäß stattfindet (oder zumindest sollte), gibt es auch den Fall dass Mitglieder anderer Teams passende Ansprechpartner darstellen können.

Der Kommunikationsfluss ist in diesem Fall jedoch nicht immer so einfach: Man hat Angst eine „dumme“ Frage zu stellen, man kennt die anderen ja vielleicht nicht so gut. Oftmals weiß man auch nicht, wen man denn fragen sollte, insbesondere wenn man nicht weiß mit was sich die Kommilitonen genau beschäftigen – wie sich die anderen Teams organisiert haben. Andersrum kann man sich auch als Antwortender unwohl fühlen, weil man befürchtet, eine „dumme“ Antwort zu geben oder fühlt sich unter Druck gesetzt irgendwas zu antworten, wenn man direkt gefragt wird.

Wie soll hier geholfen werden?

In Ihrem Praktikumssystem wurde ein Ansatz integriert, der Sie bei der Durchführung Ihres Praktikums unterstützen soll. Er bietet Ihnen die Möglichkeit Fragen anonym zu stellen – somit liegt der Fokus auf der Frage an sich, nicht auf der Person. Ebenso können Sie anonym antworten – somit liegt der Fokus auf der passenden Hilfestellung und wiederum nicht auf der Person. Der Fragende „wirft die Frage in den Raum“ ohne einen Adressaten angeben zu müssen. Das System sucht potentiell interessierte und passende Adressaten und stellt diesen die Nachricht zu. Die Empfänger können, jeder individuell, dem Fragenden helfen – ohne Ihre Identität preis zu geben. Sie können die Nachricht auch nicht beantworten. Da der Fragende nicht weiß an wen die Nachricht ging, besteht bei diesem Ansatz kein sozialer Druck zu antworten, wen man nicht will – der Fragende wird es nicht „herausfinden“.

Warum sollte ich fragen?

Während der mehr oder minder formalen Einführung in die Nutzung der SAP Systeme am Anfang des Semesters können Sie die ersten Erfahrungen sammeln. Jedoch stellt sich schnell heraus, dass der Teufel im Detail liegt. Sie werden also ganz sicher auf Herausforderungen treffen, die Sie nicht ad hoc lösen können. In diesem Fall können Sie natürlich die Dokumentation im Netz, Foren oder spezielle Bücher durchstöbern. Dies ist auch sinnvoll. Jedoch könnte man oft schneller zum Ziel kommen. Die Herausforderungen auf die man trifft sind oft allgemein genug, dass jemand anders schon eine ähnliche Lösung finden musste. Es wäre doch also gut von den Erfahrungen dieser Person profitieren zu können. „Fragen kostet nichts" heißt es im Volksmund und so ist es auch in Ihrem Praktikumssystem. Sie gehen keinerlei „Risiken" ein, da Sie anonym fragen. Das schlimmste was Ihnen passieren kann ist, dass kein passender Ansprechpartner gefunden werden konnte oder die potentiell passenden nicht antworten wollen oder können. Für Sie ist es also ein zusätzliches Medium um schneller an Ihr Ziel zu gelangen.

Warum sollte ich antworten?

Je nachdem wie tief man sich in eine Thematik hineingearbeitet kann man sich getrost als „Experte" dazu sehen. Jedoch kennt man zu „seiner" Thematik nicht immer alle Details und Facetten. Darum engagieren sich auch sehr versierte Experten in z.B. Programmierforen oder generell Online Communities. Sie lernen noch mehr, dadurch, dass Sie ähnliche Problemstellungen anderer aus einem neuen Winkel sehen. Der Lerneffekt ist hier häufig nicht zu unterschätzen, denn man erfährt auch wie andere eine Problemstellung versuchsweise angegangen sind und lernt dadurch mit. Genau das trifft bei Ihnen in verstärktem Maße zu. Wenn Sie sich in einen bestimmten Bereich im Praktikum schon mehr beschäftigt haben, profitieren Sie dennoch von Fragen anderer Teilnehmer die in etwa auf demselben Wissensstand wie Sie sein werden.

Sie „riskieren" auch als Antwortender nichts, da Sie anonym „Ihrem" Fragestellenden antworten. So kann es nicht passieren, dass jemand Sie wegen Ihrer Antwort kritisieren könnte.

Da Sie nicht in direkter Konkurrenz zu den Projekten Ihrer Kommilitonen stehen, entstehen Ihnen auch keine sonstigen Nachteile, wenn Sie sich auf eine Diskussion bzw. einen Gedankenaustausch einlassen – profitieren werden Sie aber davon.

Wie findet denn das System heraus welche Frage an wen gerichtet werden sollte?

Bekommt man eine Nachricht zu der man nicht viel sagen kann, so ist dies nicht sehr erbaulich. Je mehr dies zutrifft desto schlimmer wiegt der zumindest kurzzeitige Verlust an Aufmerksamkeit ohne dass man selbst oder der Fragende einen direkten Mehrwert hat. Auch kann man es schlicht als störend empfinden Fragen bzw. Nachrichten zu bekommen, die einen vielleicht gar nicht interessieren. Interessant sind Fragestellungen für uns am ehesten, wenn wir schon eine zumindest gewisse Erfahrung im betroffenen Themenbereich haben. Noch interessanter werden Themenstellungen die wir selbst erst kürzlich bearbeitet haben.

Diese beiden Ansätze werden in dem Implementierung, den Sie nutzen können vereint. Das System übernimmt vom Fragenden dessen Kontext, wobei Kontext hierbei bedeutet, seinen derzeitigen Arbeitsschritt im SAP System. Ausgehend von diesem Kontext sucht das System nach anderen Nutzern, die möglichst ähnlich dazu sind, also bspw. dieselbe Transaktion gerade ausführen oder erst kürzlich ausgeführt haben. Somit kann der Fragende seine Nachricht bezogen auf das was gerade tut stellen ohne dies explizit angeben zu müssen oder gar in ein anderes System wechseln zu müssen. Der Empfänger der Frage wiederum erhält nur Anfragen die seinem Tätigkeits- und Erfahrungsprofil entsprechen und ihm somit die Vertiefung seiner Erfahrungen ermöglichen.

Wie wird das System genutzt?

Der zentrale Einstiegspunkt für die Nutzung des Ansatzes geschieht über die Transaktion ZHELPME (siehe Abbildung 1).

Abbildung 1 - Startbild der Transaktion ZUSR

Hier werden alle von Ihnen im Moment genutzten Transaktionen mit Ihren Kürzeln (Transaktionscodes) sowie den sprechenden Beschreibungen angezeigt. Sie können mit einem Radiobutton auswählen zu welchem Ihrer Arbeitsschritte Sie eine Frage haben bzw. gerne Unterstützung suchen würden.

Haben Sie den gewünschten Modus gewählt und dies mit „Weiter“ bestätigt, gelangen Sie zu einer Oberfläche, die es Ihnen erlaubt eine Frage bzw. Nachricht zu verfassen (siehe Abbildung 2).

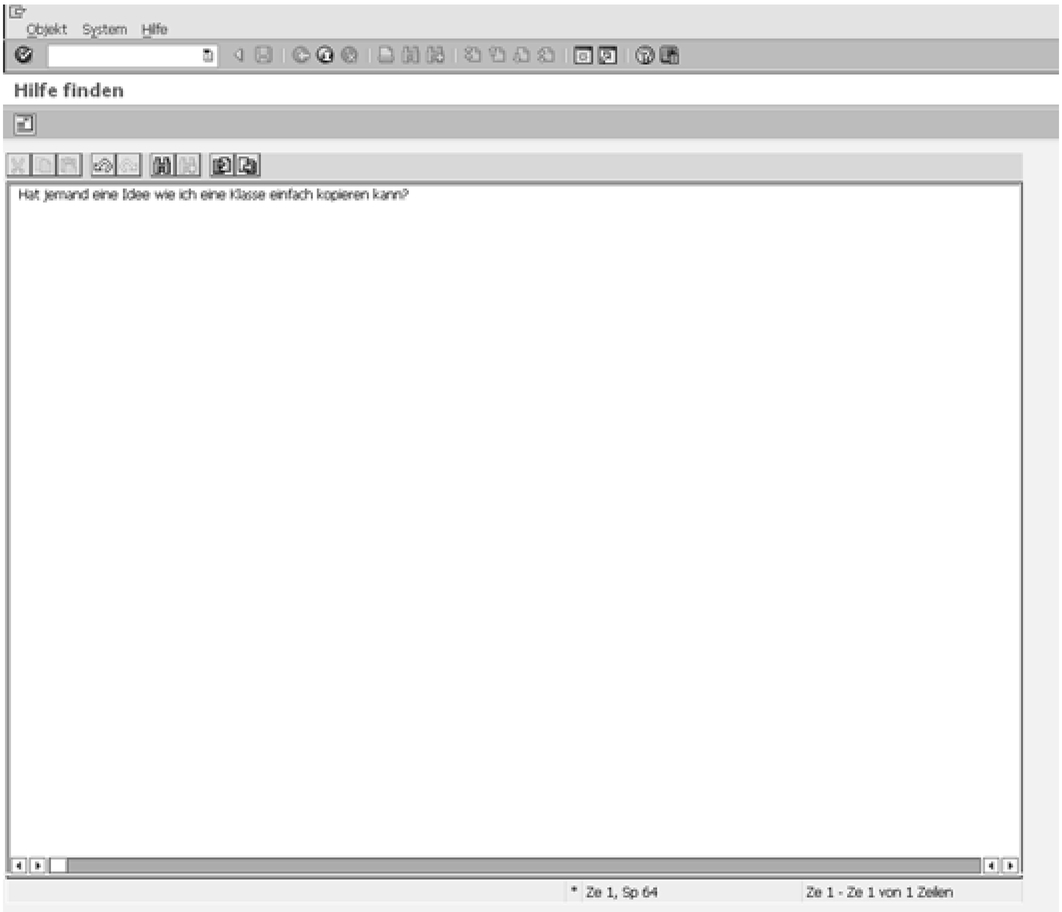


Abbildung 2 - Eingabeoberfläche für eine Frage

Sie haben hier die Möglichkeit eine Frage zu formulieren, die für Ihre Arbeit in diesem Schritt relevant ist. Sind Sie mit der Formulierung fertig, können Sie mit dem Icon das Verarbeiten und Versenden anstoßen. Die weiteren Icons können Sie für die üblichen Textbearbeitungsschritte benutzen, wenn Sie wollen. So können Sie z.B. Text ausschneiden, kopieren oder einfügen, aber auch Texte von externen Dateien einladen bzw. Ihren Text in eine externe Datei ablegen.

Haben Sie die Nachricht abgeschickt antwortet das System wie in Abbildung 3.

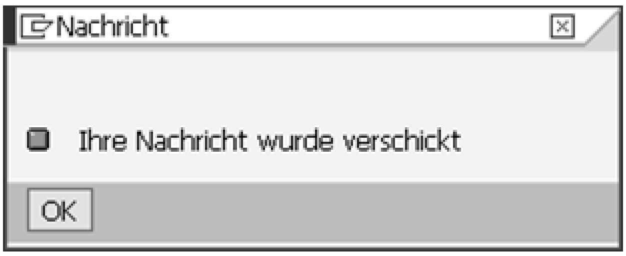


Abbildung 3 - Nachricht wurde an einen od. mehrere Empfänger verschickt

Für den Fragenden ist der Ablauf erst mal abgeschlossen. Im Hintergrund stellt das System die Nachricht nun an die passenden Nutzer zu. Falls die Nutzer gerade im System angemeldet sind, bekommen Sie eine Nachricht als Pop-Up (siehe Abbildung 4), die Sie darüber informiert, dass Ihre Expertise von einem Kollegen gefragt ist. Wie Sie schon im Screenshot sehen können, ist der Absender ein allgemeiner Nutzer, der für jegliche Kommunikation als Proxy genutzt wird.

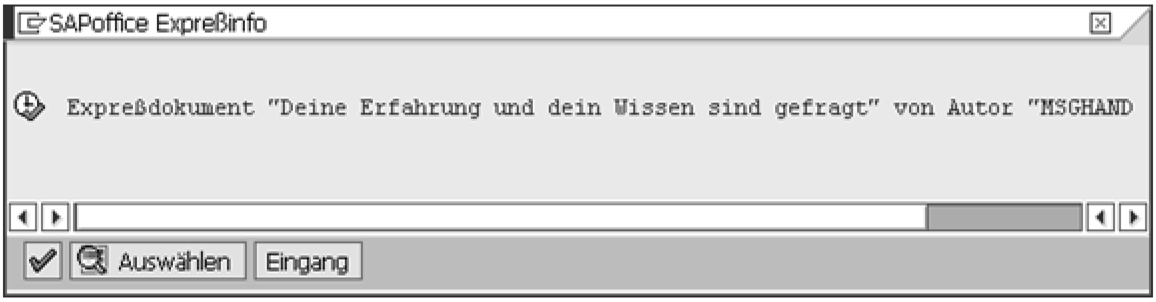


Abbildung 4 - Hinweis, dass eigene Expertise gefragt ist

Klicken Sie in diesem Pop-Up auf auf, können Sie normal weiter arbeiten. Dies ist z.B. angebracht, wenn man die Nachricht später beantworten will oder gar nicht beantworten will.

Klicken Sie hingegen auf Eingang, so gelangen Sie in Ihr persönliches Postfach. Dort finden Sie alle Nachrichten, die Ihrem Nutzer bisher zugestellt wurden. Eine Nachricht sollte wie in Abbildung 5 aussehen und Ihnen mitteilen, dass Sie jemand um Ihre Expertise zu einem Thema bittet.

Abbildung 5 - Nachricht in Persönlichem Postfach

Wenn Sie die Nachricht doppelklicken, erscheint die Mitteilung und Sie sehen welche Fragestellung an Sie herangetragen wurde (siehe Abbildung 6).

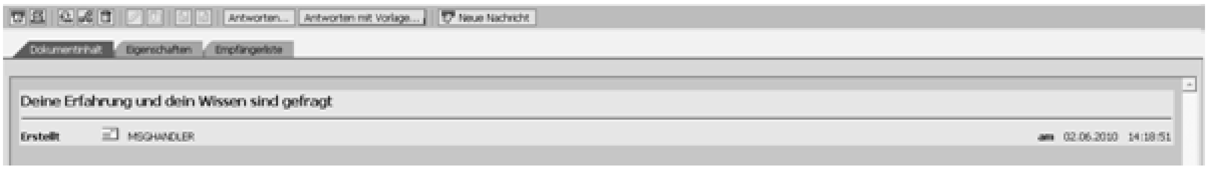


Abbildung 6 - Vollansicht einer Nachricht

Nachdem Sie die Nachricht gelesen haben, können Sie Ihre Gedanken zu der Frage weiterreichen, indem Sie auf Antworten... klicken. In der daraufhin erscheinenden Maske können Sie Ihre Antwort formulieren und diese letztlich verschicken, indem Sie auf klicken. Das System liefert Ihnen daraufhin ein Pop-Up, in dem Sie die Empfänger Ihrer Nachricht eintragen können (siehe Abbildung 7). Der zentrale Nutzer der die Verteilung Ihrer Nachricht im Hintergrund vornimmt ist bereits eingetragen, sodass Sie durch klick auf das Versenden direkt anstoßen können.

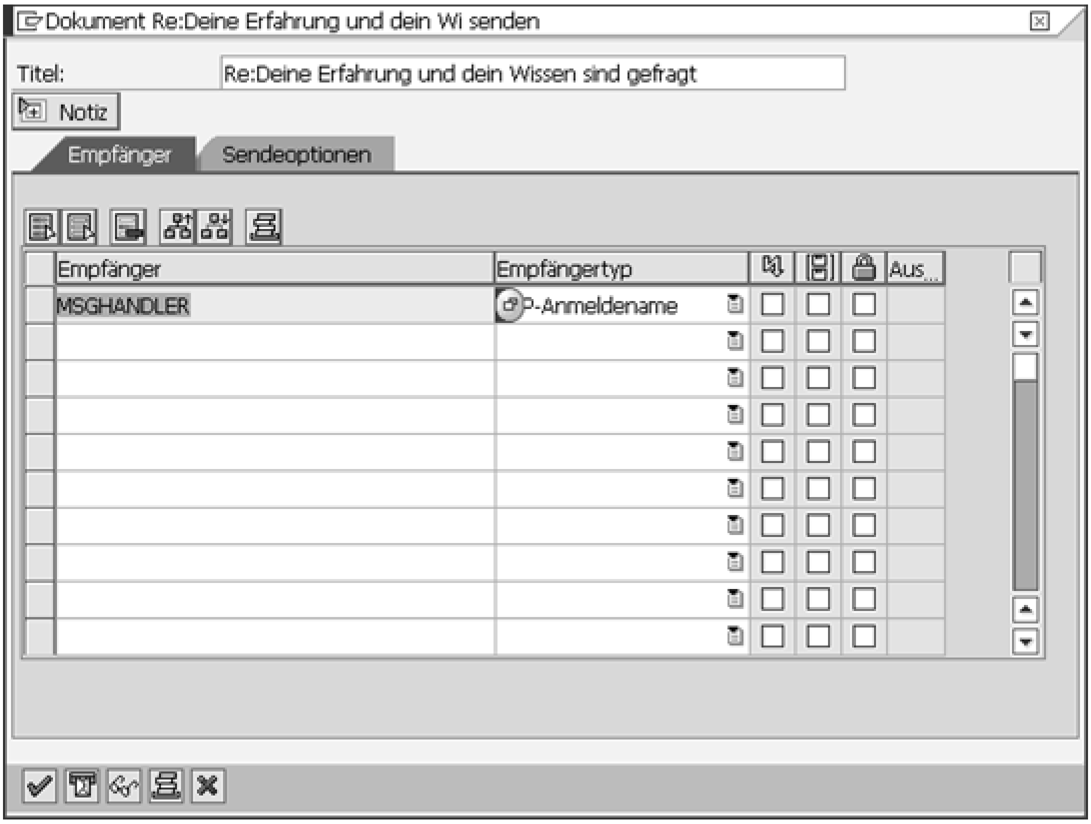


Abbildung 7 - Empfänger eintragen

Für den Antwortenden ist der Ablauf damit vorerst beendet. Die Verarbeitung der Nachricht findet in kurzen Abständen periodisch im Hintergrund statt: Dem ursprünglichen Versender der Nachricht wird Ihre Antwort zugestellt.

Sollten Sie nicht am System angemeldet sein, aber ein passender Empfänger einer Frage sein, so bekommen Sie beim nächsten Anmelden am System keine „Express-Nachricht" – also das Pop-Up, dass Sie direkt in Ihre persönliche Inbox führt. Sie gelangen jederzeit in Ihre persönliche Inbox über die Transaktion SO01 und können, z.B. gleich nach der Anmeldung am System, ihre Nachrichten überprüfen.

Hinweis: Die Kommunikation im System erfolgt anonym, d.h. In beiden Fällen sind die Sender und Empfänger einander nicht bekannt: man sieht nicht wer die Nachricht erhält und man sieht auch nicht wer geantwortet hat. Man sieht nur den Nutzer MSGHANDLER als Versender.

Weiterer Hinweis: Der dargestellte Ablauf entspricht einem Frage – Antwort Schema. Allerdings muss die Kommunikation nicht darauf beschränkt bleiben. Hat z.B. der Fragestellende noch weitere Fragen zu der Antwort eines Experten, so kann er auf die gleiche Weise, also über das Nachrichtensystem, antworten. In diesem Fall wird die Antwort auch nur dem Experten zugestellt, der geantwortet hat und nicht, wie im ersten Schritt des Ablaufs, an mehrere geeignete Experten. Dies ist nützlich um eine „Diskussion“ zwischen dem Fragenden und dem Antwortendem zu ermöglichen.

Das Verfahren klappt nicht wie beschrieben, was tun?

Sollten Sie Probleme haben, es Abbrüche geben oder das System sich nicht so verhalten wie hier beschrieben oder wie Sie es erwarten, möchte ich Sie bitten eine kurze E-Mail an joerg.schmidl@in.tum.de zu schicken.

Das System ist so schon ok, aber ich hätte gerne noch dies und jenes gesehen. Was soll ich tun?

Wenn Sie Eigenschaften am System vermissen oder sich weitere Features wünschen würden, oder wenn Ihnen Implementierungsansätze nicht gefallen und Sie der Meinung sind es würde auch besser gehen…. Ihr Feedback trifft auf mehr als offene Ohren ;-) Ich würde mich dann über eine kurze E-Mail an joerg.schmidl@in.tum.de freuen.

Warum soll ich Feedback geben bzw. mich an den Ersteller wenden?

Der implementierte Ansatz soll vornehmlich Ihnen helfen. Ob das auch gelingt, können letztlich nur die „Stakeholder“ selbst – also Sie – beurteilen. Wenn es Optimierungspotential gibt und Sie es mitteilen, wird es also zu einer Verbesserung für Sie führen können.

Der implementierte Ansatz ist mit einigem „Hirnschmalz“ entwickelt worden, wird nun aber zum ersten Mal eingesetzt. Wie auch bei Ihren eigenen Projekten, wird sich die Tragfähigkeit von Konzept und Implementierung erst bei der Anwendung zeigen. Um dies zu ermöglichen freut sich der Entwickler natürlich über Ihr Feedback und wird es Ihnen danken.

Letztlich trägt Ihr Feedback auch dazu bei, gute Forschungsergebnisse zu produzieren. Wie es das Credo der TU München und der Fakultät für Informatik im Besonderen ist, sind Sie somit aktiver Teil des wissenschaftlichen Lebens und noch viel mehr: wichtiger Bestandteil davon – Sie als Person mit Ihrem konstruktivem Feedback.

Danke vorab fürIhren Beitrag!

….und viel Erfolg bei Ihren Projekten!!!!

Jörg Schmidl

Joerg.schmidl@in.tum.de

Appendix B-2 Condensed Feedback of Target Group

Individual giving feedback	Remark	Kind of remark
Feedback 1	It is a good idea to centrally store questions that have been asked, as it reduces unnecessary double effort	Positive remarks / What I liked
Feedback 1	It is a good idea to include file upload. This allows someone to understand the problem better	Positive remarks / What I liked
Feedback 1	The overview of "questions of Transactions" allows to find out if a question or a similar one has already been posted	Positive remarks / What I liked
Feedback 1	In general the interface offers a good interface that allows to work with it easily	Positive remarks / What I liked
Feedback 1	When posting own questions they seem to not appear anywhere. It is unclear what happens to them	Negative Remark / Bug / Suggestion for improvement
Feedback 1	It would be nice to have the possibility of editing a question, or maybe answer it oneself if one finds the solution to it oneself. Maybe create it like a forum	Negative Remark / Bug / Suggestion for improvement
Feedback 1	There should be the possibility to indicate whether a question is solved or still open, e.g. using SAP's traffic lights symbols	Negative Remark / Bug / Suggestion for improvement
Feedback 2	Overview of questions according to transaction allows to see which questions have been posted already and might reduce duplicate questions	Positive remarks / What I liked
Feedback 2	Good and intuitive structure in the application that allows to easily work with it.	Positive remarks / What I liked
Feedback 2	The selection of the SAP modus (the context) allows to indicate which task the question is associated with	Positive remarks / What I liked
Feedback 2	I like the concept of having a central place for the questions (within the system). This way the affected persons can articulate their problems directly in the system very close to the moment when the question arises	Positive remarks / What I liked
Feedback 2	The possibility to upload files is good as it allows to describe questions with greater detail	Positive remarks / What I liked
Feedback 2	It is not possible to call up the application from outside	Negative Remark /

Individual giving feedback	Remark	Kind of remark
	(the main) system such as the BEx (a SAP-add-on for Excel)	Bug / Suggestion for improvement
Feedback 2	My question did not appear under myquestions nor at questions for this transaction	Negative Remark / Bug / Suggestion for improvement
Feedback 2	When a question has been posted it is unclear for the requestor when an answer is to be expected. In urgent cases his progress would be stopped. This is problematic whenever it is a critical problem	Negative Remark / Bug / Suggestion for improvement
Feedback 2	I suspect that many (of us) would still use other means of communication such as Skype because one can expect faster answers. Using the application one has to wait until someone want and does answer which might be longer	Negative Remark / Bug / Suggestion for improvement
Feedback 2	Possibly it would be better to restructure the question and answer setup. Maybe have one column for the question and another for the answers	Negative Remark / Bug / Suggestion for improvement
Feedback 3	The idea of fostering the exchange of the affected persons is generally good	Positive remarks / What I liked
Feedback 3	It is a good idea to allow for an exchange directly in the system	Positive remarks / What I liked
Feedback 3	Necessity to have a fully intuitive interface. This has been achieved almost completely already.	Negative Remark / Bug / Suggestion for improvement
Feedback 3	The application should be easy to find and quick to access in the system	Negative Remark / Bug / Suggestion for improvement
Feedback 3	Should be uncomplicated to use and quick in processing, much quicker than it is	Negative Remark / Bug / Suggestion for improvement
Feedback 3	An overview of all questions including the dates of issuing and the status should be presented to all users	Negative Remark / Bug / Suggestion for improvement
Feedback 3	It should be possible to enter a title next to the question	Negative Remark /

Individual giving feedback	**Remark**	**Kind of remark**
	itself	Bug / Suggestion for improvement
Feedback 3	It might be helpful (especially for archiving) to additionally assign a category to the questions	Negative Remark / Bug / Suggestion for improvement
Feedback 3	It would be good to have a WYSIWIG editor to have the ability to e.g. highlight portions of the question	Negative Remark / Bug / Suggestion for improvement
Feedback 3	In the "My question" tab one should see when a question was posted and the status of it, as well as potential answers to questions	Negative Remark / Bug / Suggestion for improvement
Feedback 3	An additional tab like "all questions" should show questions of all users	Negative Remark / Bug / Suggestion for improvement
Feedback 3	An incentive or reward system could be incorporated to increase the usage. Maybe the asking person can rate answers or he could assign a status like "question solved" or "still open". The one who has answered most questions in favor of the askers might be rewarded with a small bonus	Negative Remark / Bug / Suggestion for improvement
Feedback 4	Freshly posted questions seem to not appear in the my question tab	Negative Remark / Bug / Suggestion for improvement
Feedback 4	The answers to questions do not "get lost". If a new user has the same question it does not need to be answered anew	Positive remarks / What I liked
Feedback 4	It might be a good idea to have the tab with already answered questions appear first so that the likelihood of reposting a similar question is reduced. Otherwise someone might post a question and then later realize that it was answered already	Negative Remark / Bug / Suggestion for improvement
Feedback 4	In a special situation, when one is in the main menu and calls the application from there via the transaction code, the application does not pick up the context and the flow	Negative Remark / Bug / Suggestion for improvement

Individual giving feedback	Remark	Kind of remark
	of actions is broken. This bug could be fixed by showing an error message in this case	
Feedback 4	Having a headline for questions would increase the ability to quickly scan through questions	Negative Remark / Bug / Suggestion for improvement
Feedback 4	It would be beneficial if the answers were associated with the question more directly	Negative Remark / Bug / Suggestion for improvement
Feedback 4	It would be good to have a possibility to indicate if a question is still open so that is possible to quickly see where answers are available and where not	Negative Remark / Bug / Suggestion for improvement
Feedback 4	It would be better if the users did not use the SAP internal messaging system but the application for answering, possibly by linking form the message to the application. This way other solutions that were posted for the question could be seen	Negative Remark / Bug / Suggestion for improvement
Feedback 4	It would be better to answer directly in the application, which would allow to more easily clarify what a question is about	Negative Remark / Bug / Suggestion for improvement
Feedback 4	Suggestion to include a newsfeed that is dependent on the parts of the SAP menu that one uses	Negative Remark / Bug / Suggestion for improvement
Feedback 4	On a newsfeed page one should see related questions – one's own and the ones of others – as well as the answers of other subscribers. In essence this could be a forum filtered according to subscriptions. Whoever subscribed automatically gets answers to open questions.	Negative Remark / Bug / Suggestion for improvement
Feedback 5	It is possible to ask anonymously	Positive remarks / What I liked
Feedback 5	The application is directly integrated into the system	Positive remarks / What I liked
Feedback 5	There is SAP's official help site and forums where it is possible to ask experts	Negative Remark / Bug / Suggestion

Individual giving feedback	**Remark**	**Kind of remark**
		for improvement
Feedback 5	It would be good to have the possibility to choose if a question is to be posted anonymously or with one's user name	Negative Remark / Bug / Suggestion for improvement
Feedback 5	Anonymous questions give rise to the risk of exploiting the application	Negative Remark / Bug / Suggestion for improvement
Feedback 5	It would be beneficial to select experts or groups of individuals to whom one's question is sent. The other users are then not disturbed	Negative Remark / Bug / Suggestion for improvement
Feedback 5	The use of tabs renders the application easy to use	Positive remarks / What I liked
Feedback 5	Responding to questions is easy, the interface is comfortable to use	Positive remarks / What I liked
Feedback 5	In the "My question" tab, it is tedious to have to click on refresh before one sees the question. It should be displayed immediately	Negative Remark / Bug / Suggestion for improvement
Feedback 5	There is no documentation for the application. Users would like to know how to work with the transaction maybe illustrated with an example	Negative Remark / Bug / Suggestion for improvement
Feedback 5	Some terminology like "modus" are not explained.	Negative Remark / Bug / Suggestion for improvement
Feedback 5	It is not obvious in which transactions the application is available. Maybe it would be possible to integrate an icon there	Negative Remark / Bug / Suggestion for improvement
Feedback 5	It appears to be only possible to open the application in the main menu. In other transactions it is necessary to open a new window which is extra effort	Negative Remark / Bug / Suggestion for improvement
Feedback 5	It appears to be impossible to deactivate participation in the help function. What if I do not want to receive questions	Negative Remark / Bug / Suggestion for improvement
Feedback 5	It would be nice to have the possibility to create a screen-	Negative Remark /

Individual giving feedback	Remark	Kind of remark
	shot and integrate it directly into a question to better explain the problem	Bug / Suggestion for improvement
Feedback 5	(De-)activation of the functionality and (de-) activation of anonymity should be possible	Negative Remark / Bug / Suggestion for improvement
Feedback 5	It is not possible to delete questions that have been posted once, which however is sometimes necessary	Negative Remark / Bug / Suggestion for improvement
Feedback 5	In a special situation, when one is in the main menu and calls the application from there via the transaction code, the application does not pick up the context and the flow of actions is broken. This bug could be fixed by showing an error message in this case	Negative Remark / Bug / Suggestion for improvement
Feedback 5	There is a bug, when one closes the answer pop-up without entering a text, the application is closed, not just the pop-up	Negative Remark / Bug / Suggestion for improvement

Table A-1 ***Feedback to first iteration application***
Source: Own illustration

Appendix C Documents Used and Created in Second Case Study

Appendix C-1 Initial Instruction Notes Delivered to Target Group

E-Mail sent to participants

Hallo liebe SAPiens,

ich möchte euch auf ein neues Support Tool aufmerksam machen, dass euch ab sofort (sobald das System wieder erreichbar ist) zur Verfügung steht. Dabei handelt es sich um ein Student-zu-Student Support und soll vor allem die Kommunikation und Interaktion zwischen Euch stärken. Der große Reiz bei dem neuen Tool besteht darin, dass ihr es direkt aus dem SAP System aufruft und alle Interaktion auch im SAP System stattfindet. Im Anhang findet Ihr auch eine Beschreibung des neuen Tools und wie es verwendet werden kann.

Welche Vorteile bringt mir das Stellen einer Frage in der Anwendung:

- Ich kann direkt aus SAP eine Frage stellen, ohne in ein anderes Programm wechseln zu müssen und der Kontext der Frage wird mitgenutzt
- Ich sehe Fragen, die andere Teilnehmer haben und lerne so von Ihnen
- Die Frage geht an die richtigen Leute - diejenigen die schon mal was im gleichen Kontext gemacht haben
- Die Frage kann sich zu einer ausgedehnten Diskussion erweitern. So kann man deutlich mehr lernen
- Fragen in der Anwendung werden auch vom Kurssupport (Leitung und Mentoren) "gesehen" und können von diesen direkt in der Anwendung bearbeitet werden. Weil der Kontext ersichtlich ist und ggf. schon Diskussion statt gefunden hat auch deutlich genauer.

Welche Vorteile bringt es mir eine Frage, die mir zugestellt wird zu beantworten:

- Ich kann als erfahrener Teilnehmer mein bisheriges Wissen erweitern und vernetzen, indem ich die Fragestellungen anderer mit ihnen und weiteren bespreche
- Die Fragen die ich bekomme, sind mir zugestellt, weil ich in dem Kontext schon gearbeitet hab. Daher sind die "Fragenden" gute Diskussionspartner
- Antworte ich auf Fragen, bin ich in die weitere Diskussionen "eingeklinkt" und bin schneller Up-to-date. Ich bekomme mehr mit und lerne mehr durch die Diskussion
- Wenn vom Teilnehmer gewünscht, wird sein (frei wählbarer) Name statt einem anonymen Kürzel angezeigt. Somit sehen andere Nutzer wer sich in bestimmten Bereichen besonders gut auskennt und sich in Diskussionen einbringen konnte.
- Wenn vom Teilnehmer gewünscht, wird er in die Liste der aktivsten Teilnehmer aufgenommen, die für alle sichtbar ist und herausstellt, wer sich durch sein besonderes Wissen hervorheben konnte.

- Der eigene Beitrag wird auch entlohnt. Die Sechs aktivsten Teilnehmer bekommen am Ende des Kurses SAP-Bücher als Anerkennung Ihrer Leistung.

Ich werde natürlich auch auf Anfragen antworten (falls diese nicht zeitig gelöst werden), aber werde vor allen Dingen Euch selbst die Gelegenheit geben, Probleme zu erörtern und zu lösen.

Ich hoffe, auf eine rege Beteiligung und würde mich auch über Feedback zu dem neuen Tool freuen.

Viele Grüße
XXXXXX
(project coordinator – name anonymized)

Attached descriptive document in PDF format (order: left to right, then up down)

Hilfesuche und Expertenvermittlung in SAP Anwendungen

Gegenseitige Hilfe für Studenten in SAP-basierten Lernszenarien

Motivation für die Anwendung

Hintergrund, Kontext und Zielvorstellung

Ausgangssituation und Problemstellung

- In SAP-basierten Lehrveranstaltungen ist die Betreuung der Studenten sehr intensiv
  - Nutzung des Interfaces und seine Eigenheiten muss von den Studenten erlernt werden
  - SAP-spezifische Umsetzung von Geschäftsprozessen muss von den Studenten verinnerlicht werden
  - Betriebswirtschaftlicher bzw. technischer Hintergrund (je nach Zielgruppe) muss allgemein verstanden werden

Lösungsansätze I

- Typische Hilfequellen für den Studenten und ihre Nachteile
  - SAP Help: meist nur allgemeine Einführung in Konzepte, wenig Varianz hinsichtlich verschiedener Alternativen
  - SAP Service Market Place: Für Studenten i.A. nicht zugreifbar, wenig spezifische Information
  - SAP OSS: Wenig kontextualisiert auf Anwendungskontext, Fokus auf technischen Problemen, nicht Anwendungswissen

Lösungsansätze II

- Weitere typische Hilfequellen für den Studenten und ihre Nachteile
  – WWW: Sehr unstrukturiert; Aufwand für Suche hoch
  – Social Networks & Foren (z.B. SDN): Teils sehr viele Einträge -> Suche zeitintensiv weil nicht kontextualisiert, Medienbruch bei Frage in SAP System
  – Selbsthilfe: Zeitaufwendig, Gefahr des „Reinventing the wheel“ Phänomens

Lösungsansätze III

- Und noch eine typische Hilfequellen für den Studenten….

…Kursbetreuer und Tutoren

Probleme…
– Skaliert nicht so gut
– Kann sehr aufwendig, zeitintensiv und damit teuer werden

Weiterer didaktischer Aspekt – Wie lernen Studenten…

- Lernen aus formalen Unterlagen (printed oder online oder im Frontal“unterricht“)
- Durchführen von strukturierte Aufgaben (z.B. im Rahmen von eLearnings; detaillierte Fallstudien)
- Unstrukturierte Aufgaben (meist im Rahmen von Praktika; Ziel wird vorgegeben)
- Aber auch und meist am Besten: Beschäftigen miteigenen Problemen und denen anderer.

Idealvorstellung

Hilfe zur „Gruppenselbsthilfe“

Was sind also erstrebenswerte Ziele I

- Studenten sollen sich gegenseitig helfen um mit und von einander zu lernen und so komplexe Sachverhalte besser durchdringen zu können
- Betreuer und Veranstalter sollen soweit es möglich ist entlastet werden
- Medienbrüche sollen vermieden werden damit Hilfe-suche und Diskussion zu Fragen direkt im SAP System statt finden kann

Was sind also erstrebenswerte Ziele II

- Fragen / Diskussionen sollen „konserviert“ werden können um „Reinventing the wheel“ zu vermeiden
- Fragen / Diskussionen sollen für die Studenten verständlich strukturiert und auffindbar sein idealerweise nach Arbeitskontext
- Da sich Studenten nicht zwingend (alle) kennen, soll auffinden passender „Experten“ keine Aufgabe des Fragenden sein

Die Lösung

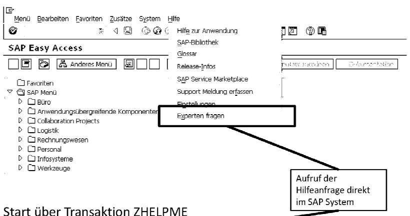
- Eine Anwendung die..
  …direkt im SAP System implementiert ist und von dort aufrufbar ist
  …Fragen des Nutzers aus seinem Kontext heraus an-nimmt und nach dem Kontext (=Transaktion) strukturiert
  …eine gute Balance zwischen zuviel und zuwenig Information bietet
  …die Zuordnung zwischen Fragenden und „Experten“ automatisch übernimmt

Nutzung der Anwendung

Überblick

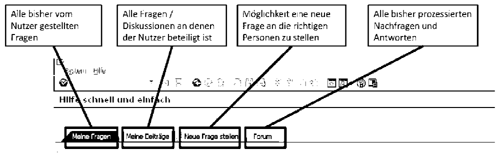
Starten der Anwendung

a) Start über Hilfemenü



b) Start über Transaktion ZHELPME

Übersicht über die Anwendung



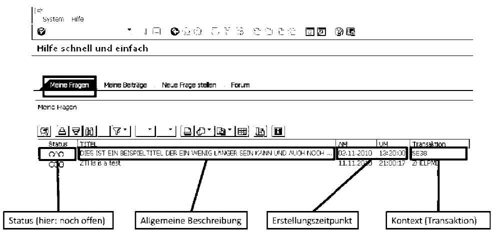
Nutzung der Anwendung

Ansicht *Meine Fragen*

Beschreibung der Ansicht *Meine Fragen*

- Nutzer soll die Fragen sehen, die er gestellt hat
- Fragen können gelöst sein oder noch offen (Status)
- Fragen haben einen knappen Titel zur allgemeinen Beschreibung (Titel)
- Der zeitliche Kontext einer Frage kann eine Rolle spielen (Zeitpunkt)
- Der Arbeitskontext einer Frage ist relevant (Transaktion)

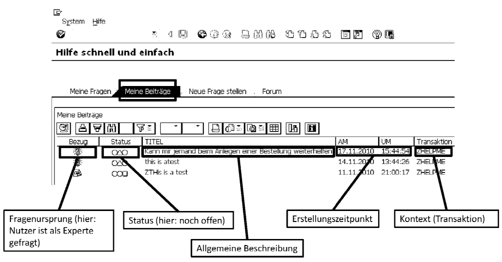
Ansicht *Meine Fragen*



Beschreibung der Ansicht *Meine Beiträge*

- Nutzer soll Fragen sehen, die für ihn/sie relevant sind
- Relevant sind Fragen, wenn
  a) der Nutzer als Experte für die Frage gilt (Bezug 1)
  b) wenn er unabhängig davon vorher auf eine Frage geantwortet hat (Bezug 2)
- Fragen können gelöst sein oder noch offen (Status)
- Fragen haben einen knappen Titel zur allgemeinen Beschreibung (Titel)
- Der zeitliche Kontext einer Frage kann eine Rolle spielen (Zeitpunkt)
- Der Arbeitskontext einer Frage ist relevant (Transaktion)

Ansicht *Meine Beiträge*



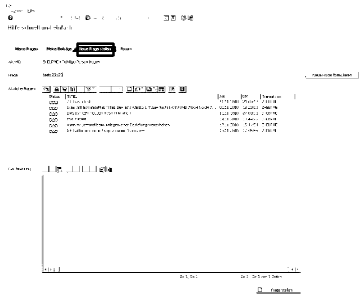
Nutzung der Anwendung

Ansicht *Neue Frage stellen*

Beschreibung der Ansicht *Neue Frage stellen*

- Nutzer soll eine allgemeine Beschreibung eingeben können (Frage)
- Der Nutzer soll einen Fragenkontext spezifizieren können (Aktivität/Modus)
- Der Nutzer soll eine detaillierte Beschreibung eingeben können (Beschreibung)
- Der Nutzer soll zur Wiederverwendung von Wissen angeregt werden (Ähnliche Fragen)

Ansicht *Neue Frage stellen* (Gesamtsicht)



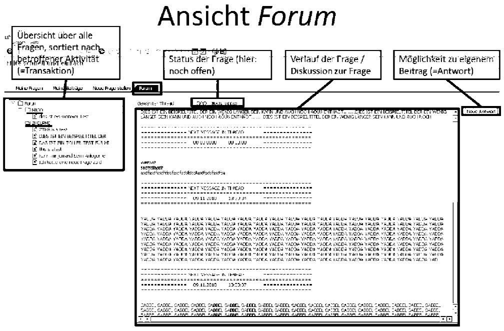
Nutzung der Anwendung

Ansicht *Forum*

Beschreibung der Ansicht *Forum*

- Nutzer soll zusätzlich, bei weiterem Interesse in allen Fragen und Diskussionen suchen können um Wissen wiederzuverwenden
- Nutzer soll Verlauf sehen können um aus diesem zusätzlich lernen zu können
- Nutzer soll sich in beliebigen Frageverläufen einbringen können
- Die Fragen und Diskussionen sollen nach relevanten Aktivitäten (=Transaktionen) strukturiert sein

Ansicht *Forum*



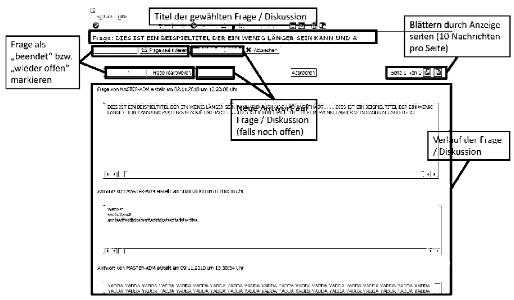
Nutzung der Anwendung

Ansicht *Nachrichtenverlauf*

Beschreibung der Ansicht *Nachrichtenverlauf*

- Nutzer soll gesamten Verlauf der Frage sehen
- Wenn der Nutzer die Frage selbst gestellt hat, soll er sie „schließen" können (=gelöst)
- Falls die Frage eines Nutzers geschlossen ist, aber doch wieder relevant wird, soll sie wieder reaktiviert werden können
- Wenn weitere Informationen zu der Frage beigetragen werden sollen, soll der Nutzer eine neue Antwort abgeben können
- Potentiell viele Beiträge zur Frage sollen verwaltbar sein

Ansicht Nachrichtenverlauf (Aufruf aus meine Fragen / meine Beteiligungen)



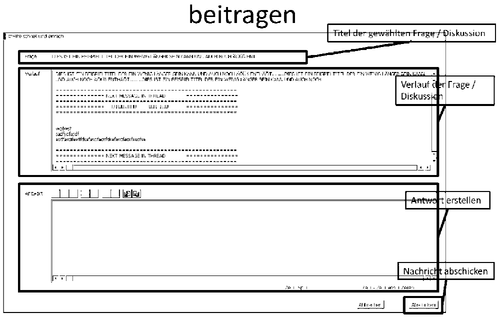
Nutzung der Anwendung

Ansicht *Antworten*

Beschreibung der Ansicht *Antworten*

- Nutzer soll zu einem Beitrag den Titel sehen können (Frage)
- Nutzer soll zu einem Beitrag den gesamten bisherigen Verlauf sehen können (Verlauf)
- Die Antworteingabe soll durch einen (SAP-) Editor unterstützt werden (Antwort)

Antworten auf Fragen / Zur Diskussion beitragen



Nutzung der Anwendung

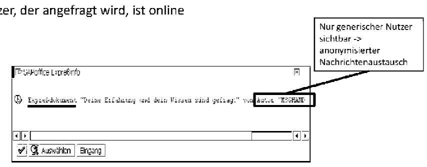
Benachrichtigung der Nutzer über Interaktion

Beschreibung der Benachrichtigung

- Nutzer, der automatisch als Experte bestimmt wurde, soll von neuer Anfrage informiert werden
- Um Medienbruch zu vermeiden soll das SAP-interne Nachrichtensystem genutzt werden
- Wenn der „Experte“ gerade online ist, soll er sofort benachrichtigt werden (Expressdokument)

Benachrichtigung bei Expertenanfrage / Änderung in beteiligter Diskussion

a) Nutzer, der angefragt wird, ist online



b) Start über Transaktion ZHELPME

→Hinweis auf neue Nachricht in SAP Inbox (TA: SBWP) des Nutzers

Ergebnis und Vorteile der Anwendung

Vorteile

- Studenten können sich gegenseitig helfen und dadurch von einander lernen
- Durch die Verankerung im SAP System gibt es keinen Medienbruch
- Die Sortierung nach Transaktionen erlaubt es den Studenten, nach Wissen in Ihrem derzeitigen Arbeitskontext zu suchen
- Die Vermittlung von Experten entbindet den Studenten von der Schwierigkeit zu wissen, „wer helfen kann“

Appendix C-2 Updated Instruction Notes Delivered to Target Group

E-Mail sent to participants

Liebe Sapiens,

das Kursende rückt näher und der Endspurt beginnt :-)

Im Laufe des Semesters hatten wir euch eine SAP-interne Hilfemöglichkeit (Transaktion ZHELPME) angeboten. Um euch in den letzten Wochen weiterhin optimal unterstützen zu können, wurde heute eine neue Version der Anwendung eingespielt, die einige Verbesserungen und Erweiterungen mit sich bringt. Damit ihr damit zurecht kommt, hab ich euch auch die neueste Version der Dokumentation dazu angehängt. In diesem Zuge auch noch der Hinweis: Mitmachen lohnt sich! Die sechs Studenten die unterstützt von der Anwendung am meisten Expertise zeigen, dürfen sich über ein SAP Press Buch freuen. Grüße

XXXX
(project coordinator – anonymized)

Attached descriptive document in PDF format (order: from left to right, then up down)

Hilfesuche und Expertenvermittlung in SAP Anwendungen

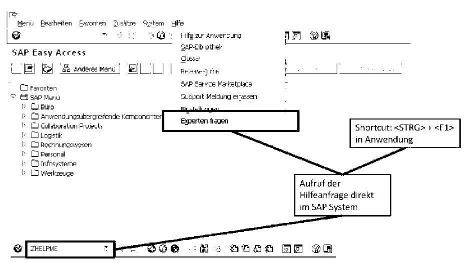
Gegenseitige Hilfe für Studenten in SAP-basierten Lernszenarien

Vorteile

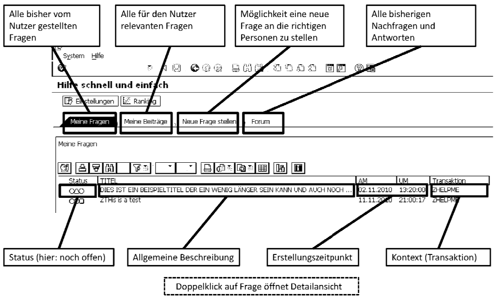
- Studenten können sich gegenseitig helfen und dadurch von einander lernen
- Durch die Verankerung im SAP System gibt es keinen Medienbruch, kein Abspringen in andere Anwendungen / das Web ist mehr nötig
- Die Sortierung nach Transaktionen erlaubt das Suchen nach Wissen im derzeitigen Arbeitskontext
- Die Vermittlung von Experten entbindet den Studenten von der Schwierigkeit zu wissen, „wer helfen kann“
- Die 6 Studenten, die Ihre Expertise am meisten unter Beweis stellen konnten, dürfen sich über SAP-Bücher freuen

Handhabung und Funktion der Anwendung

Starten der Anwendung



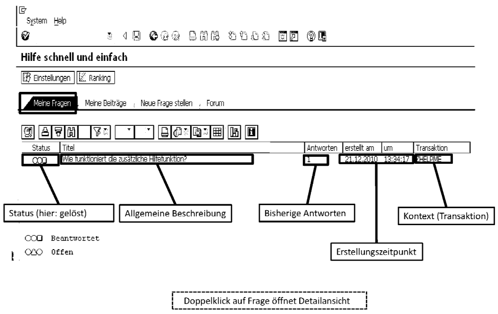
Übersicht über die Anwendung



Ansicht *Meine Fragen*

- Hier sind alle Fragen und deren Antwortverläufe zu finden, die von mir selbst gestartet wurden – bei denen **ich** also als erster was **gefragt habe**

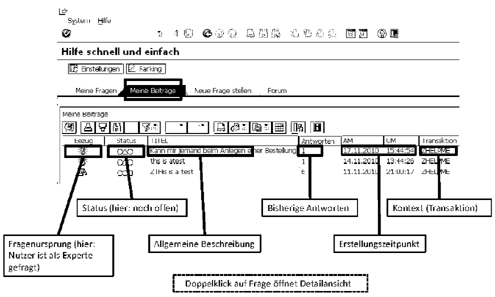
Ansicht Meine Fragen



Ansicht *Meine Beiträge*

- Hier sind alle Fragen und deren Antwortverläufe zu finden, bei…
  - …denen **ich als Experte** von jemanden um Unterstützung gebeten wurde
  - …denen ich mich in die Diskussion **selbst eingebracht** habe (aus dem Forum heraus)

Ansicht *Meine Beiträge*

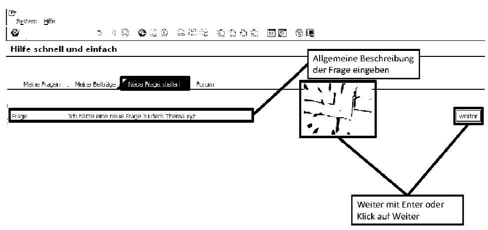


Neue Frage stellen

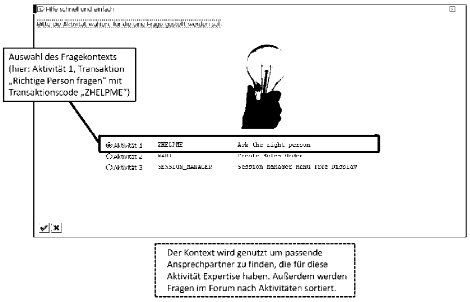
- Hier kann ich eine Frage direkt im SAP System stellen. Durch die Zuordnung einer Aktivität (= Transaktion in SAP) wird die Frage und folgende Diskussion archiviert. Außerdem nutzt die Anwendung die Aktivität um die Frage an passende „Sapiens" zu schicken

Ansicht *Neue Frage stellen* (Gesamtsicht)

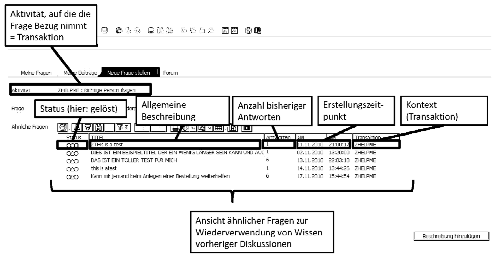
Schritt 1: Titel der Frage eingeben



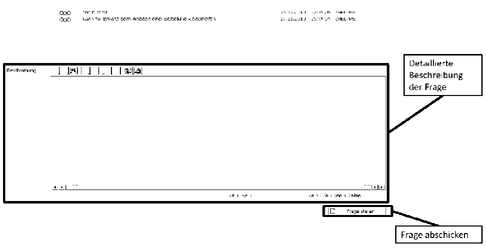
Schritt 2: Kontext (=Aktivität) der Frage auswählen



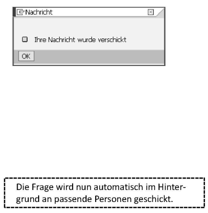
Schritt 3: Fragen mit gleicher Aktivität überblicken



Schritt 4: Detaillierte Beschreibung der Frage eingeben



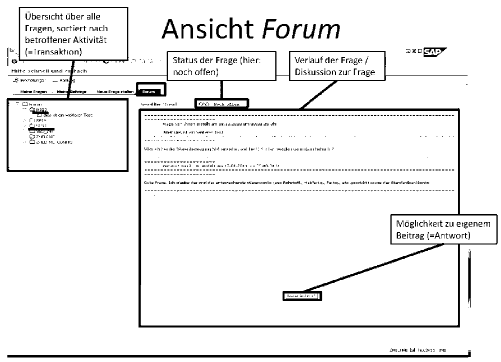
Schritt 5: Frage ist verschickt



Forum

- Das Forum enthält alle bisher gestellten Fragen und die dazugehörigen Diskussionen. Die Fragen sind thematisch nach den Aktivitäten (= Transaktionen in SAP) geordnet zu denen die Frage gestellt wurde. Ein Indikator zeigt an, ob eine Frage noch „offen" ist. Wenn ja und wenn ich mitdiskutieren kann und will, ist vom Forum aus eine Möglichkeit hierzu gegeben. Die Diskussion erscheint in diesem Fall zusätzlich in „Meine Beiträge".

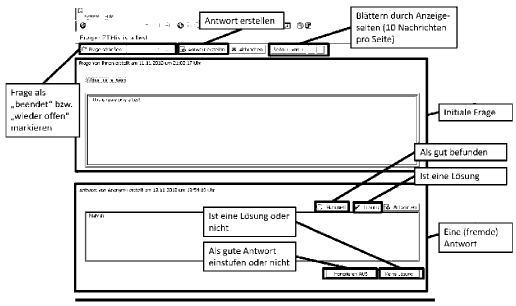
Ansicht *Forum*



Detailansicht Nachricht (Aufruf aus meine Fragen / meine Beteiligungen)

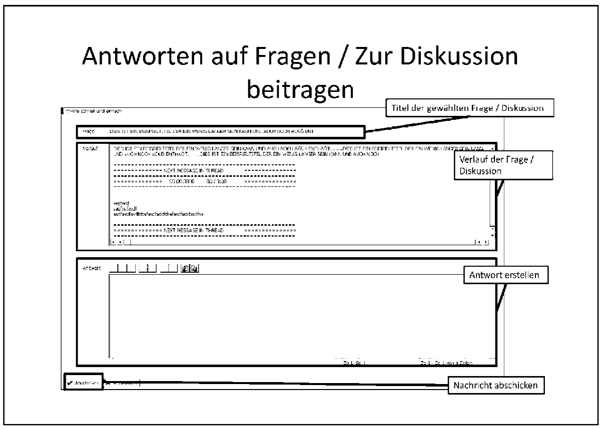
- Frage schließen: Keine weiteren Antworten mehr möglich. Diskussion ist aus Sicht des Fragenden abgeschlossen
- Honorieren: Einzelne Antworten können honoriert werden -> Feedback an den Antwortgebenden
- Lösung: Einzelne Antworten können die Frage (komplett oder zu großen Teilen) lösen und werden dann als Lösung markiert

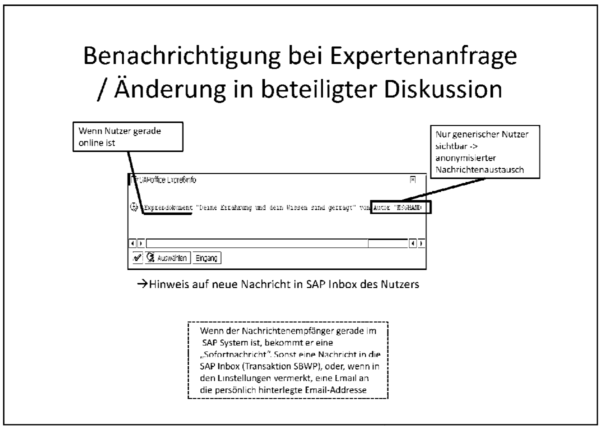
Detailansicht Nachricht (Aufruf aus meine Fragen / meine Beteiligungen)

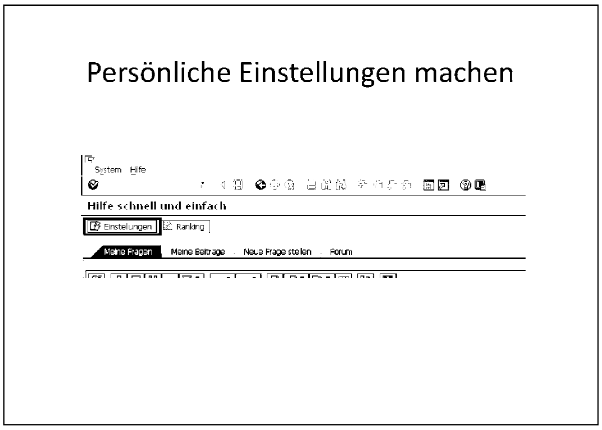


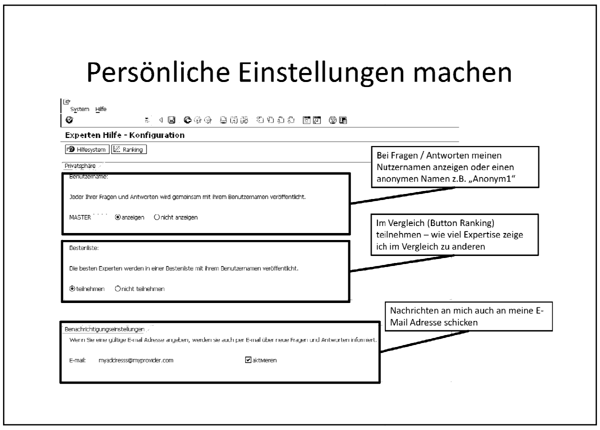
Antworten auf Fragen / Zur Diskussion beitragen

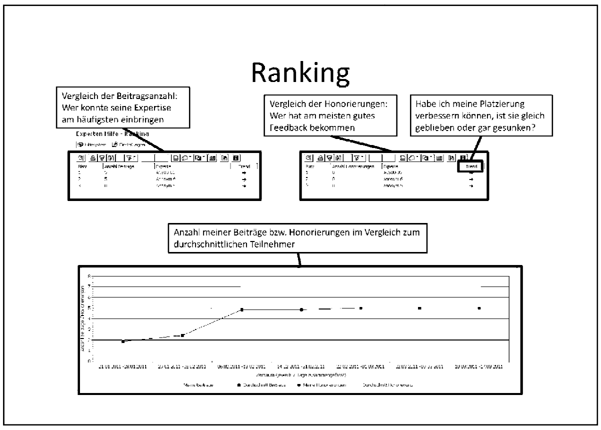
- Antworten werden automatisch an den Fragenden und alle bisher Beteiligten zugestellt











Viel Erfolg beim SAP lernen!

Appendix C-3 Online Survey Distributed to Participants of Second Case Study

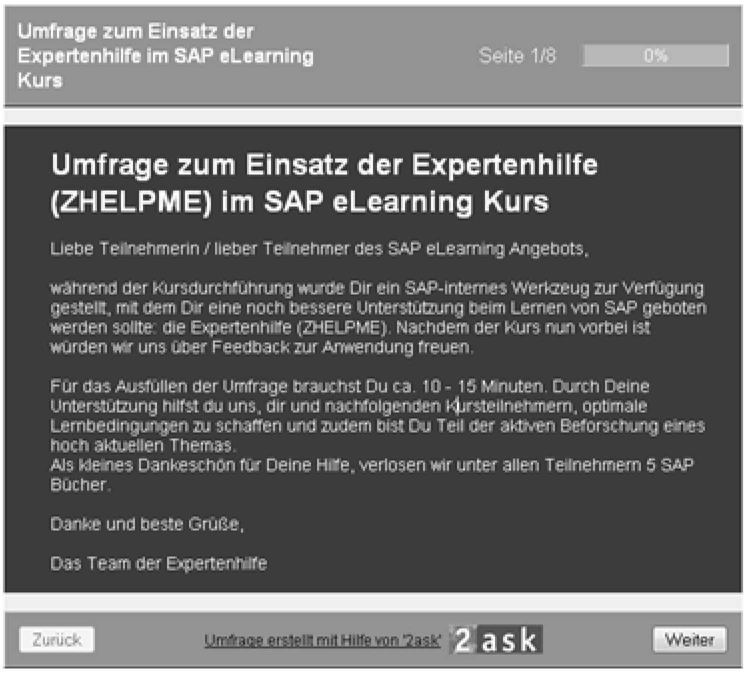


Figure C-1 ***First screen of online survey distributed to the participants of the second case study***
Source: Own illustration

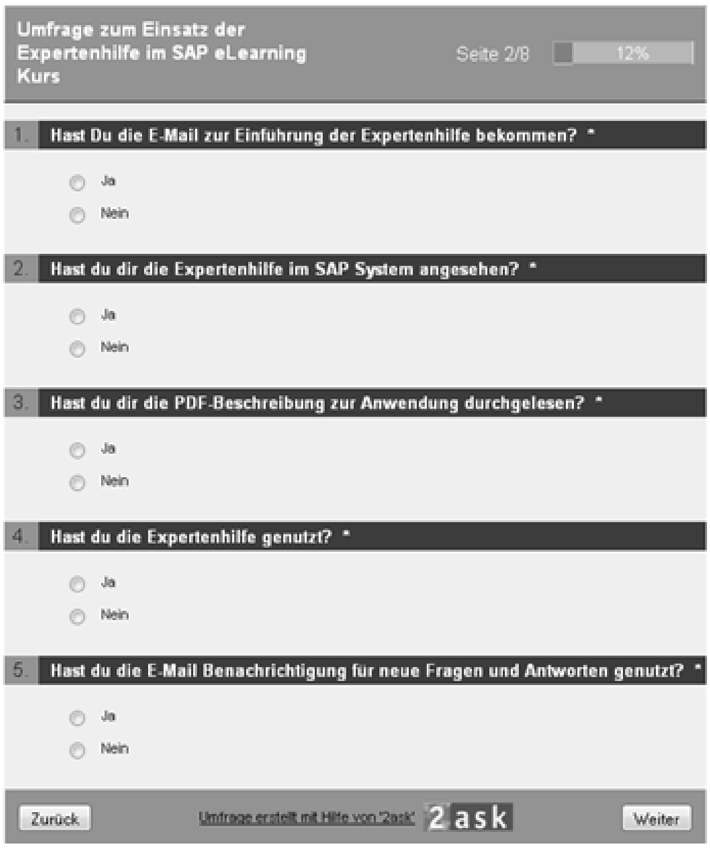
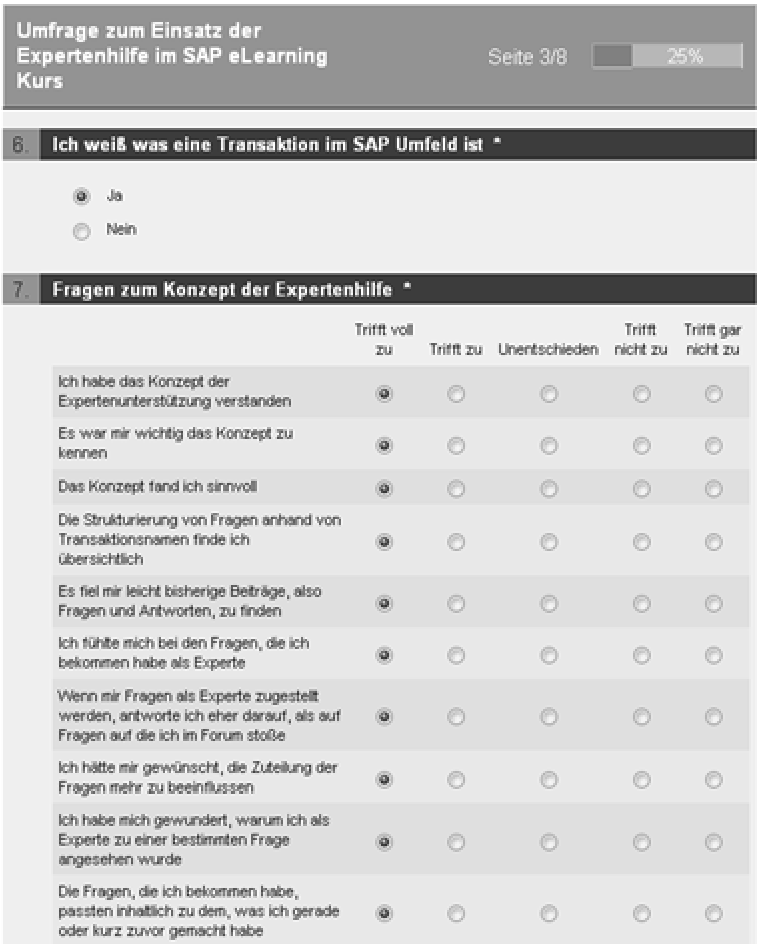


Figure C-2 ***Second screen of online survey distributed to the participants of the second case study***
Source: Own illustration



Umfrage zum Einsatz der Expertenhilfe im SAP eLearning Kurs

Seite 3/8 25%

6. Ich weiß was eine Transaktion im SAP Umfeld ist *

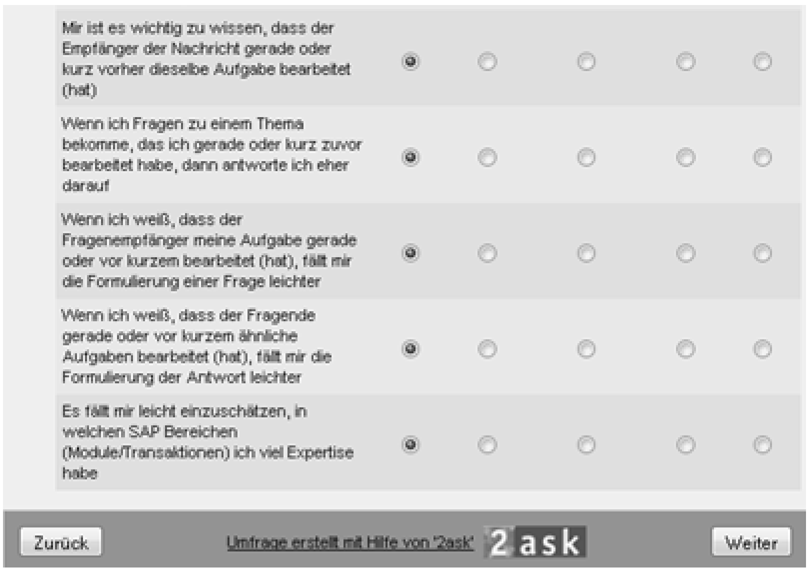
Ja

Nein

7. Fragen zum Konzept der Expertenhilfe *

	Trifft voll zu	Trifft zu	Unentschieden	Trifft nicht zu	Trifft gar nicht zu
Ich habe das Konzept der Expertenunterstützung verstanden					
Es war mir wichtig das Konzept zu kennen					
Das Konzept fand ich sinnvoll					
Die Strukturierung von Fragen anhand von Transaktionsnamen finde ich übersichtlich					
Es fiel mir leicht bisherige Beiträge, also Fragen und Antworten, zu finden					
Ich fühlte mich bei den Fragen, die ich bekommen habe als Experte					
Wenn mir Fragen als Experte zugestellt werden, antworte ich eher darauf, als auf Fragen auf die ich im Forum stoße					
Ich hätte mir gewünscht, die Zuteilung der Fragen mehr zu beeinflussen					
Ich habe mich gewundert, warum ich als Experte zu einer bestimmten Frage angesehen wurde					
Die Fragen, die ich bekommen habe, passten inhaltlich zu dem, was ich gerade oder kurz zuvor gemacht habe					

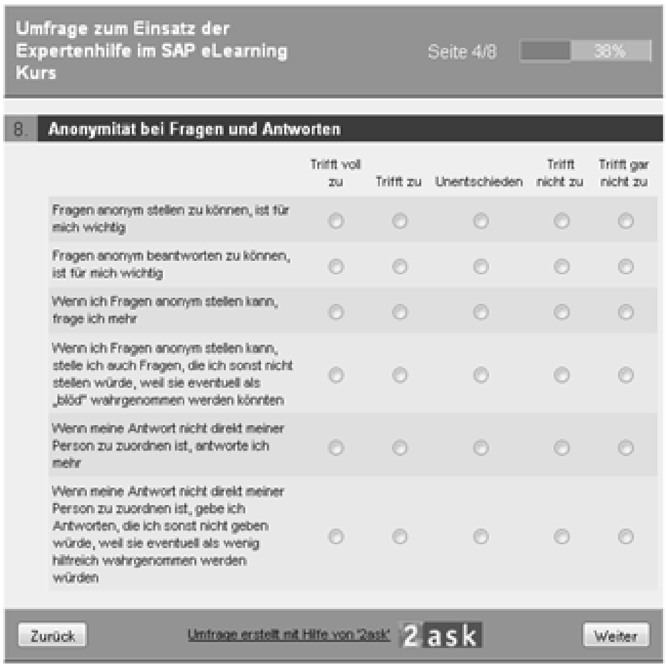
Figure C-3 ***Third screen of online survey distributed to the participants of the second case study- first part***
Source: Own illustration



Mir ist es wichtig zu wissen, dass der Empfänger der Nachricht gerade oder kurz vorher dieselbe Aufgabe bearbeitet (hat)					
Wenn ich Fragen zu einem Thema bekomme, das ich gerade oder kurz zuvor bearbeitet habe, dann antworte ich eher darauf					
Wenn ich weiß, dass der Fragenempfänger meine Aufgabe gerade oder vor kurzem bearbeitet (hat), fällt mir die Formulierung einer Frage leichter					
Wenn ich weiß, dass der Fragende gerade oder vor kurzem ähnliche Aufgaben bearbeitet (hat), fällt mir die Formulierung der Antwort leichter					
Es fällt mir leicht einzuschätzen, in welchen SAP Bereichen (Module/Transaktionen) ich viel Expertise habe					

Zurück Umfrage erstellt mit Hilfe von '2ask' 2ask Weiter

Figure C-4 ***Third screen of online survey distributed to the participants of the second case study- second part***
Source: Own illustration

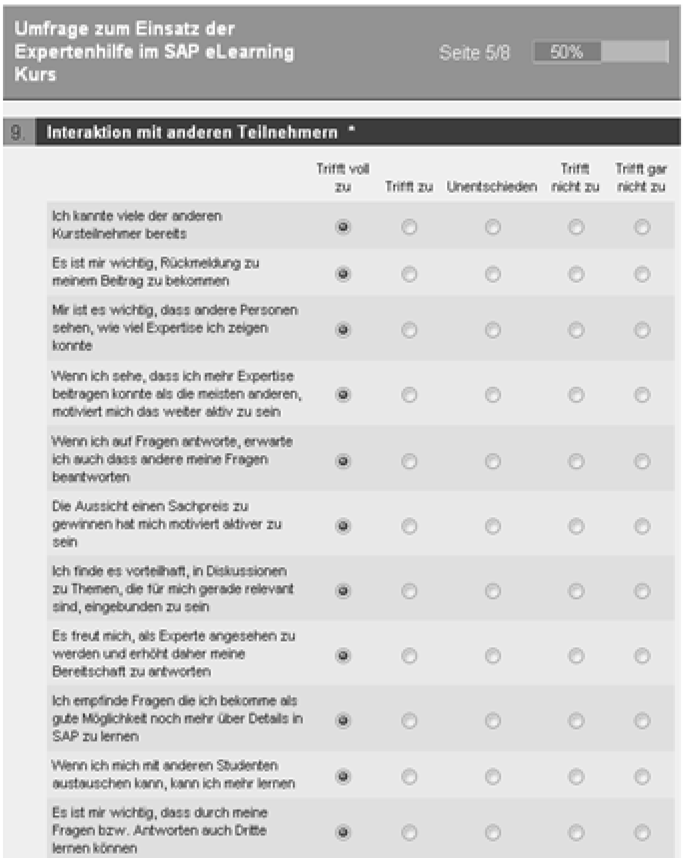

Umfrage zum Einsatz der Expertenhilfe im SAP eLearning Kurs
Seite 4/8 38%

8. Anonymität bei Fragen und Antworten

	Trifft voll zu	Trifft zu	Unentschieden	Trifft nicht zu	Trifft gar nicht zu
Fragen anonym stellen zu können, ist für mich wichtig					
Fragen anonym beantworten zu können, ist für mich wichtig					
Wenn ich Fragen anonym stellen kann, frage ich mehr					
Wenn ich Fragen anonym stellen kann, stelle ich auch Fragen, die ich sonst nicht stellen würde, weil sie eventuell als „blöd" wahrgenommen werden könnten					
Wenn meine Antwort nicht direkt meiner Person zu zuordnen ist, antworte ich mehr					
Wenn meine Antwort nicht direkt meiner Person zu zuordnen ist, gebe ich Antworten, die ich sonst nicht geben würde, weil sie eventuell als wenig hilfreich wahrgenommen werden würden					

Zurück Umfrage erstellt mit Hilfe von '2ask' 2ask Weiter

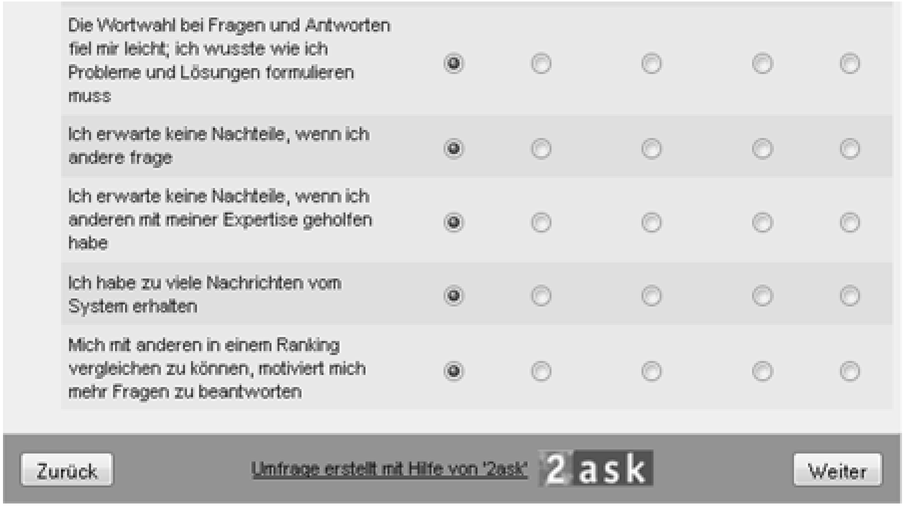
Figure C-5 ***Fourth screen of online survey distributed to the participants of the second case study***
Source: Own illustration


Umfrage zum Einsatz der Expertenhilfe im SAP eLearning Kurs
Seite 5/8 50%

9. Interaktion mit anderen Teilnehmern *

	Trifft voll zu	Trifft zu	Unentschieden	Trifft nicht zu	Trifft gar nicht zu
Ich kannte viele der anderen Kursteilnehmer bereits	●				
Es ist mir wichtig, Rückmeldung zu meinem Beitrag zu bekommen	●				
Mir ist es wichtig, dass andere Personen sehen, wie viel Expertise ich zeigen konnte	●				
Wenn ich sehe, dass ich mehr Expertise beitragen konnte als die meisten anderen, motiviert mich das weiter aktiv zu sein	●				
Wenn ich auf Fragen antworte, erwarte ich auch dass andere meine Fragen beantworten	●				
Die Aussicht einen Sachpreis zu gewinnen hat mich motiviert aktiver zu sein	●				
Ich finde es vorteilhaft, in Diskussionen zu Themen, die für mich gerade relevant sind, eingebunden zu sein	●				
Es freut mich, als Experte angesehen zu werden und erhöht daher meine Bereitschaft zu antworten	●				
Ich empfinde Fragen die ich bekomme als gute Möglichkeit noch mehr über Details in SAP zu lernen	●				
Wenn ich mich mit anderen Studenten austauschen kann, kann ich mehr lernen	●				
Es ist mir wichtig, dass durch meine Fragen bzw. Antworten auch Dritte lernen können	●				

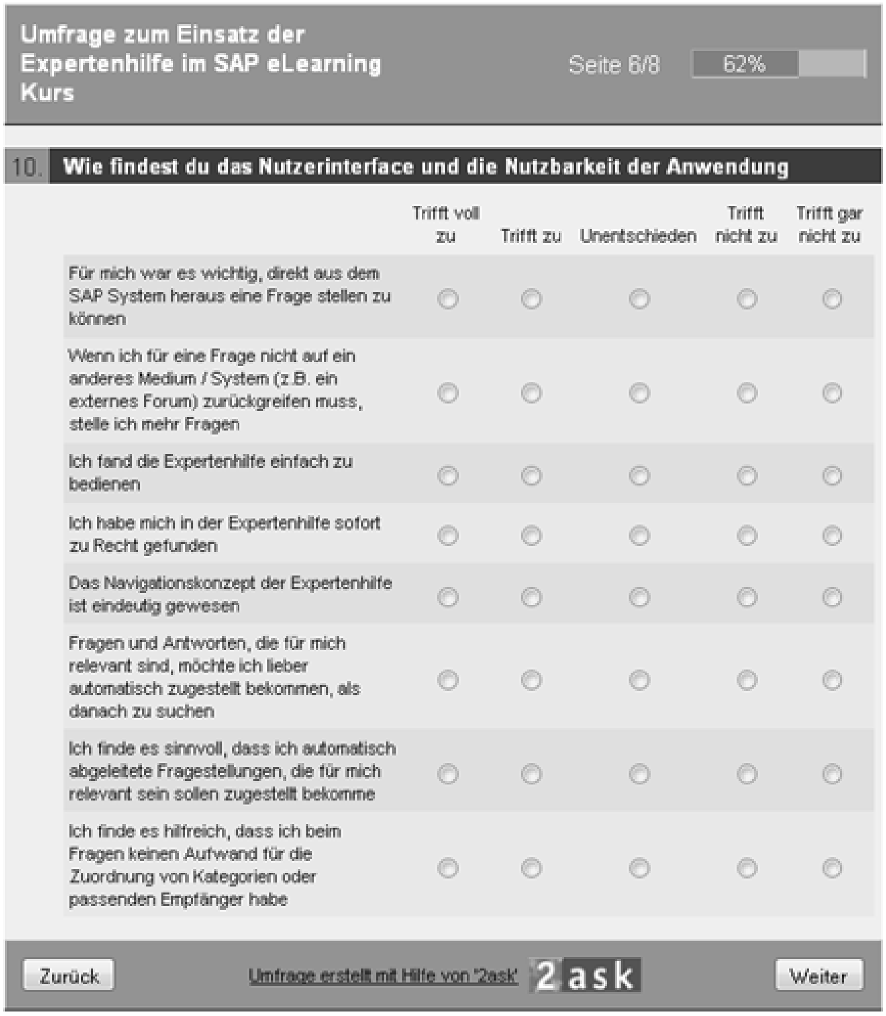
Figure C-6 ***Fifth screen of online survey distributed to the participants of the second case study – first part***
Source: Own illustration



Die Wortwahl bei Fragen und Antworten fiel mir leicht; ich wusste wie ich Probleme und Lösungen formulieren muss	◉	○	○	○	○
Ich erwarte keine Nachteile, wenn ich andere frage	◉	○	○	○	○
Ich erwarte keine Nachteile, wenn ich anderen mit meiner Expertise geholfen habe	◉	○	○	○	○
Ich habe zu viele Nachrichten vom System erhalten	◉	○	○	○	○
Mich mit anderen in einem Ranking vergleichen zu können, motiviert mich mehr Fragen zu beantworten	◉	○	○	○	○

Zurück | Umfrage erstellt mit Hilfe von '2ask' 2ask | Weiter

Figure C-7 ***Fifth screen of online survey distributed to the participants of the second case study – second part***
Source: Own illustration



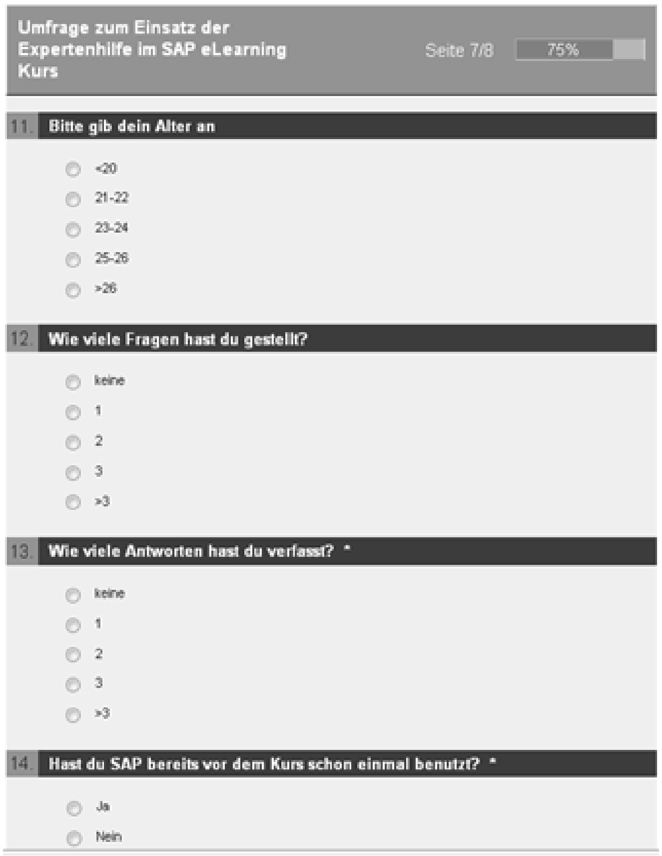
Umfrage zum Einsatz der Expertenhilfe im SAP eLearning Kurs Seite 6/8 62%

10. Wie findest du das Nutzerinterface und die Nutzbarkeit der Anwendung

	Trifft voll zu	Trifft zu	Unentschieden	Trifft nicht zu	Trifft gar nicht zu
Für mich war es wichtig, direkt aus dem SAP System heraus eine Frage stellen zu können	○	○	○	○	○
Wenn ich für eine Frage nicht auf ein anderes Medium / System (z.B. ein externes Forum) zurückgreifen muss, stelle ich mehr Fragen	○	○	○	○	○
Ich fand die Expertenhilfe einfach zu bedienen	○	○	○	○	○
Ich habe mich in der Expertenhilfe sofort zu Recht gefunden	○	○	○	○	○
Das Navigationskonzept der Expertenhilfe ist eindeutig gewesen	○	○	○	○	○
Fragen und Antworten, die für mich relevant sind, möchte ich lieber automatisch zugestellt bekommen, als danach zu suchen	○	○	○	○	○
Ich finde es sinnvoll, dass ich automatisch abgeleitete Fragestellungen, die für mich relevant sein sollen zugestellt bekomme	○	○	○	○	○
Ich finde es hilfreich, dass ich beim Fragen keinen Aufwand für die Zuordnung von Kategorien oder passenden Empfänger habe	○	○	○	○	○

Zurück | Umfrage erstellt mit Hilfe von '2ask' 2ask | Weiter

Figure C-8 ***Sixth screen of online survey distributed to the participants of the second case study***
Source: Own illustration



Umfrage zum Einsatz der Expertenhilfe im SAP eLearning Kurs

Seite 7/8 75%

11. Bitte gib dein Alter an

- <20
- 21-22
- 23-24
- 25-26
- >26

12. Wie viele Fragen hast du gestellt?

- keine
- 1
- 2
- 3
- >3

13. Wie viele Antworten hast du verfasst? *

- keine
- 1
- 2
- 3
- >3

14. Hast du SAP bereits vor dem Kurs schon einmal benutzt? *

- Ja
- Nein

Figure C-9 ***Seventh screen of online survey distributed to the participants of the second case study – first part***
Source: Own illustration

15. Ich habe vor diesem SAP Kurs bereits andere SAP Kurse besucht / an anderen SAP Kursen teilgenommen *

- nein
- ja, und zwar an [Anzahl Kurse]

16. Internetforen nutze ich... *

- sehr häufig
- häufig
- hin und wieder
- selten
- nie

17. Wenn ich Fragen habe nutze ich bisher normalerweise zuerst das folgende Medium *

- Instant Messaging (z.B. ICQ)
- E-Mail
- das Kurs-eigene Forum
- öffentliches Forum (z.B. SAP Developer Network)
- Persönliche Treffen
- Telefonkonferenzen (z.B. via Skype)
- Sonstiges:

Figure C-10 ***Second screen of online survey distributed to the participants of the second case study – second part***
Source: Own illustration

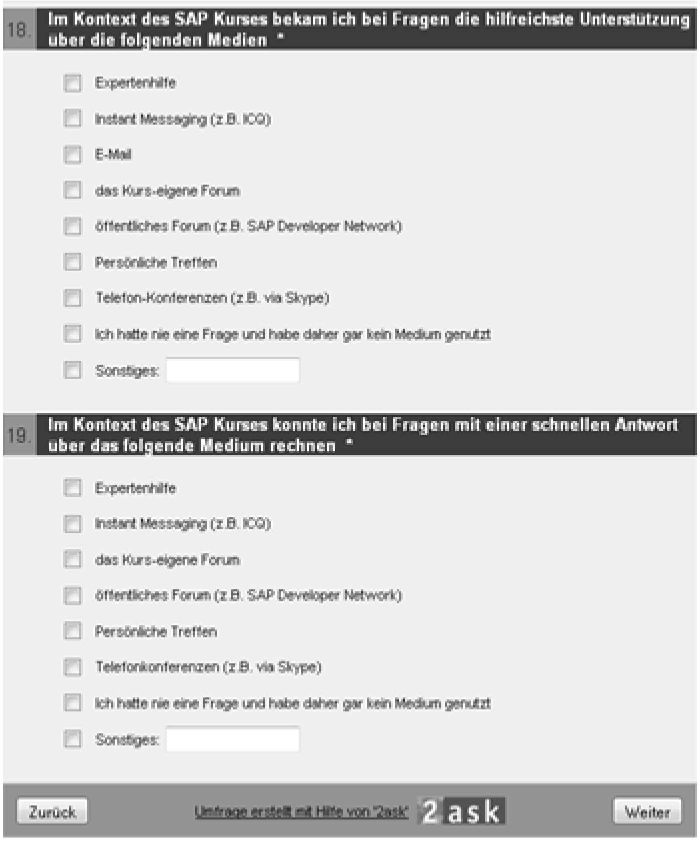
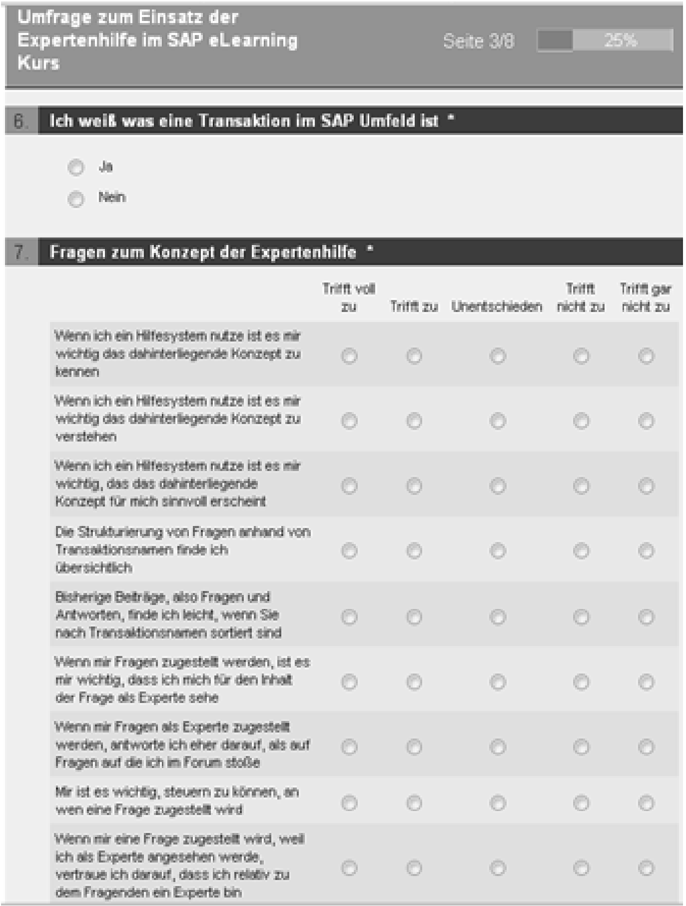


Figure C-11 ***Second screen of online survey distributed to the participants of the second case study – third part***
Source: Own illustration

Figure C-12 ***Eighth screen of online survey distributed to the participants of the second case***
Source: Own illustration



Umfrage zum Einsatz der Expertenhilfe im SAP eLearning Kurs — Seite 3/8 — 25%

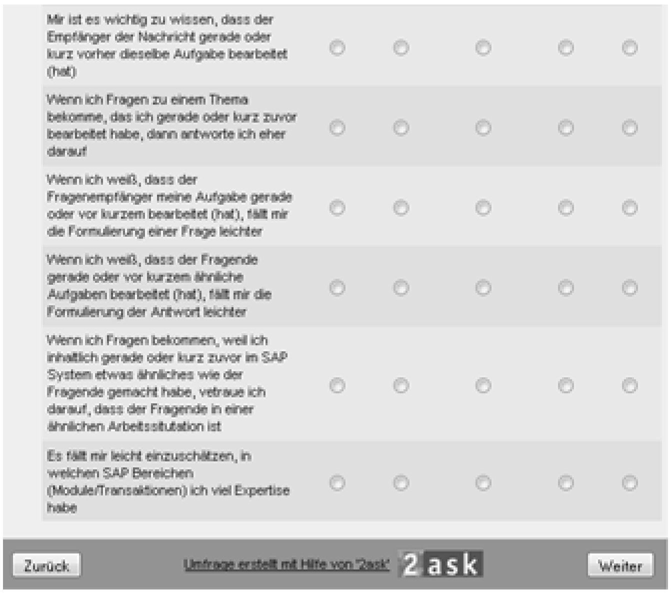
6. Ich weiß was eine Transaktion im SAP Umfeld ist *

- Ja
- Nein

7. Fragen zum Konzept der Expertenhilfe *

	Trifft voll zu	Trifft zu	Unentschieden	Trifft nicht zu	Trifft gar nicht zu
Wenn ich ein Hilfesystem nutze ist es mir wichtig das dahinterliegende Konzept zu kennen	○	○	○	○	○
Wenn ich ein Hilfesystem nutze ist es mir wichtig das dahinterliegende Konzept zu verstehen	○	○	○	○	○
Wenn ich ein Hilfesystem nutze ist es mir wichtig, das das dahinterliegende Konzept für mich sinnvoll erscheint	○	○	○	○	○
Die Strukturierung von Fragen anhand von Transaktionsnamen finde ich übersichtlich	○	○	○	○	○
Bisherige Beiträge, also Fragen und Antworten, finde ich leicht, wenn Sie nach Transaktionsnamen sortiert sind	○	○	○	○	○
Wenn mir Fragen zugestellt werden, ist es mir wichtig, dass ich mich für den Inhalt der Frage als Experte sehe	○	○	○	○	○
Wenn mir Fragen als Experte zugestellt werden, antworte ich eher darauf, als auf Fragen auf die ich im Forum stoße	○	○	○	○	○
Mir ist es wichtig, steuern zu können, an wen eine Frage zugestellt wird	○	○	○	○	○
Wenn mir eine Frage zugestellt wird, weil ich als Experte angesehen werde, vertraue ich darauf, dass ich relativ zu dem Fragenden ein Experte bin	○	○	○	○	○

Figure C-13 *Alternative third screen of online survey distributed to the participants of the second case, visible to participants that did not have any expose to this thesis' approach- first part*
Source: Own illustration



Mir ist es wichtig zu wissen, dass der Empfänger der Nachricht gerade oder kurz vorher dieselbe Aufgabe bearbeitet (hat)	○	○	○	○	○
Wenn ich Fragen zu einem Thema bekomme, das ich gerade oder kurz zuvor bearbeitet habe, dann antworte ich eher darauf	○	○	○	○	○
Wenn ich weiß, dass der Fragenempfänger meine Aufgabe gerade oder vor kurzem bearbeitet (hat), fällt mir die Formulierung einer Frage leichter	○	○	○	○	○
Wenn ich weiß, dass der Fragende gerade oder vor kurzem ähnliche Aufgaben bearbeitet (hat), fällt mir die Formulierung der Antwort leichter	○	○	○	○	○
Wenn ich Fragen bekommen, weil ich inhaltlich gerade oder kurz zuvor im SAP System etwas ähnliches wie der Fragende gemacht habe, vetraue ich darauf, dass der Fragende in einer ähnlichen Arbeitssitutation ist	○	○	○	○	○
Es fällt mir leicht einzuschätzen, in welchen SAP Bereichen (Module/Transaktionen) ich viel Expertise habe	○	○	○	○	○

Zurück — Umfrage erstellt mit Hilfe von '2ask' 2ask — Weiter

Figure C-14 *Alternative third screen of online survey distributed to the participants of the second case, visible to participants that did not have any expose to this thesis' approach- second part*
Source: Own illustration

Appendix C-4 Results of the Second Case Study's Online Survey

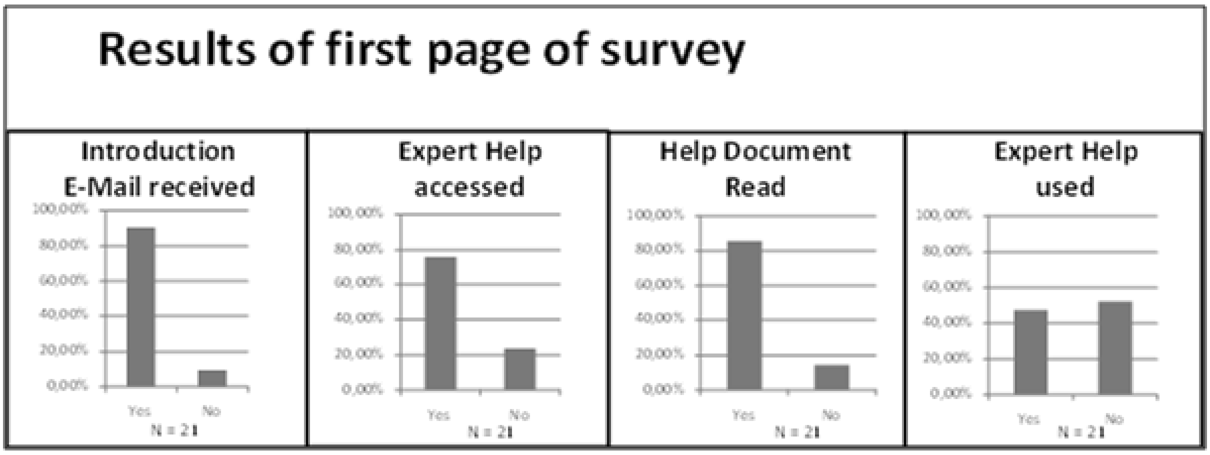


Figure C-15 ***Results of the second case study's survey – first survey page***
Source: Own illustration

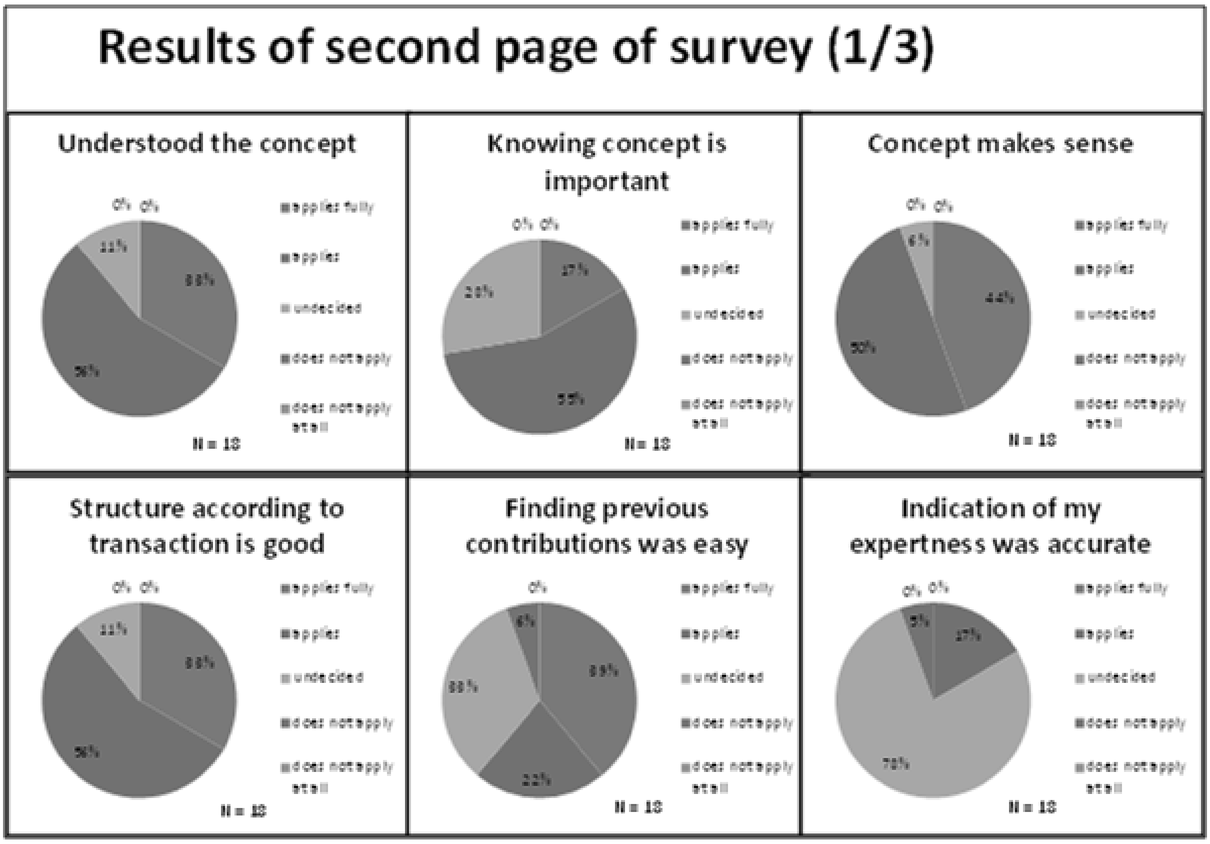


Figure C-16 ***Results of the second case study's survey – second survey page, part 1 of 3***
Source: Own illustration

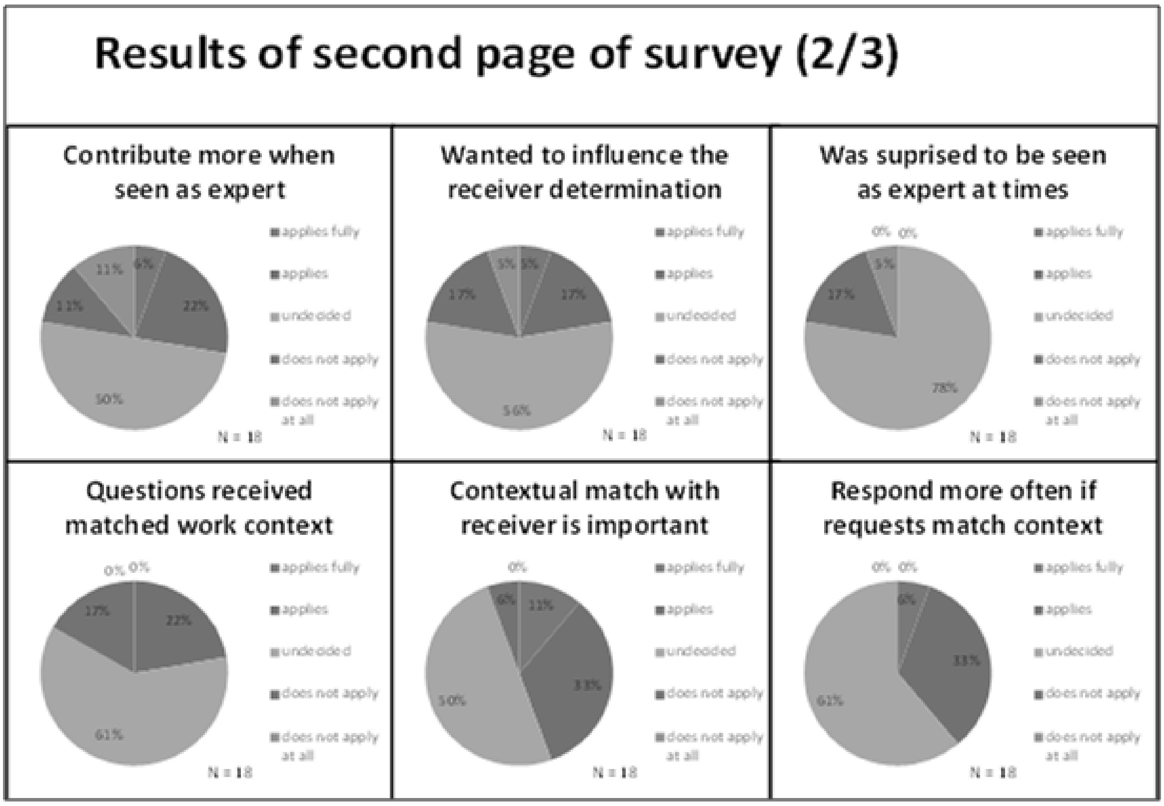


Figure C-17 ***Results of the second case study's survey – second survey page, part 2 of 3***
Source: Own illustration

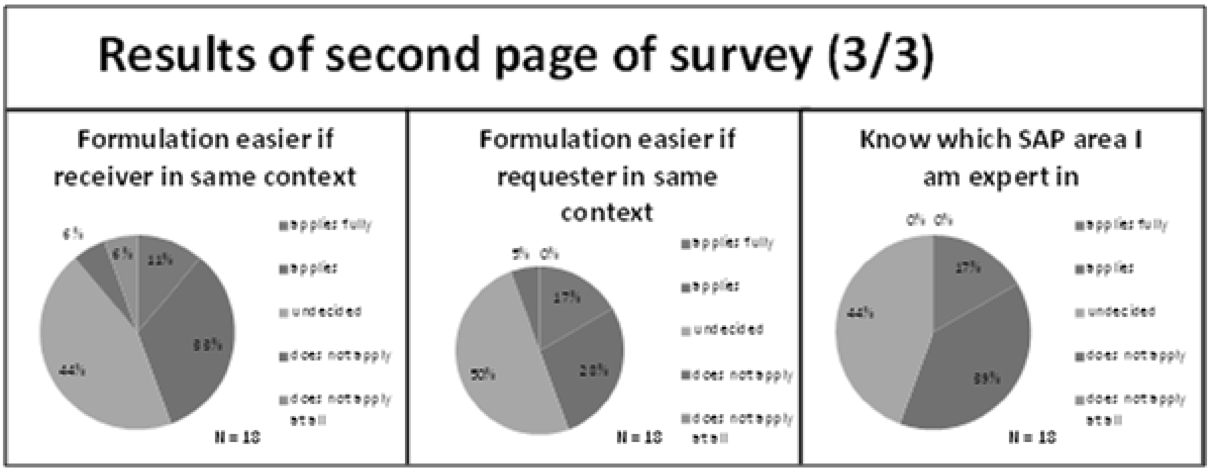


Figure C-18 ***Results of the second case study's survey – second survey page, part 3 of 3***
Source: Own illustration

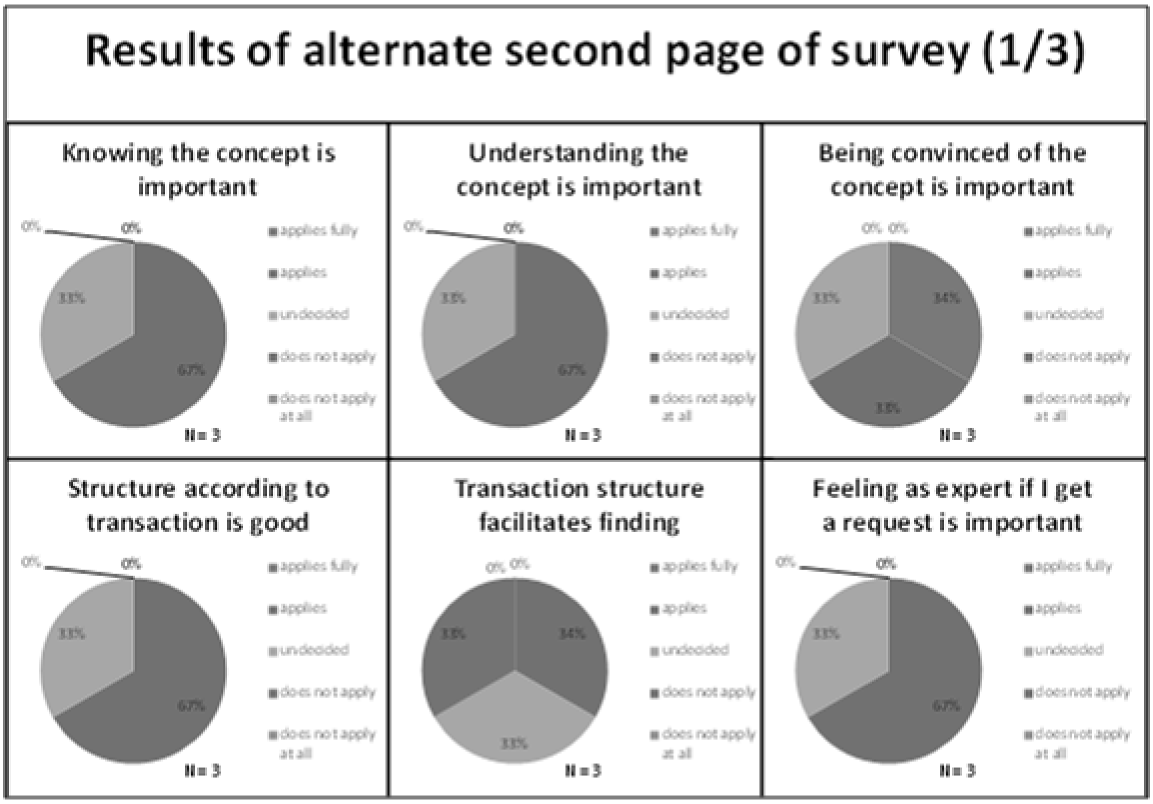


Figure C-19 ***Results of the second case study's survey – alternate second survey page, part 1 of 3***
Source: Own illustration

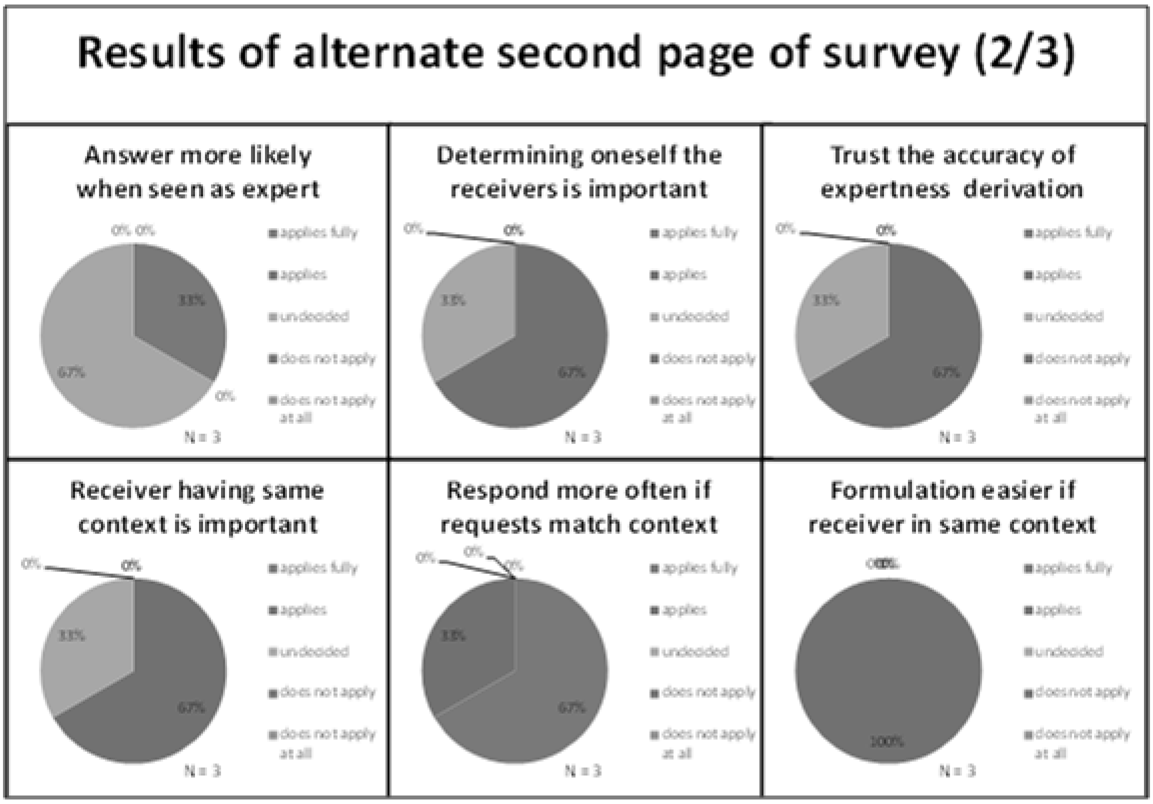


Figure C-20 ***Results of the second case study's survey – alternate second survey page, part 2 of 3***
Source: Own illustration

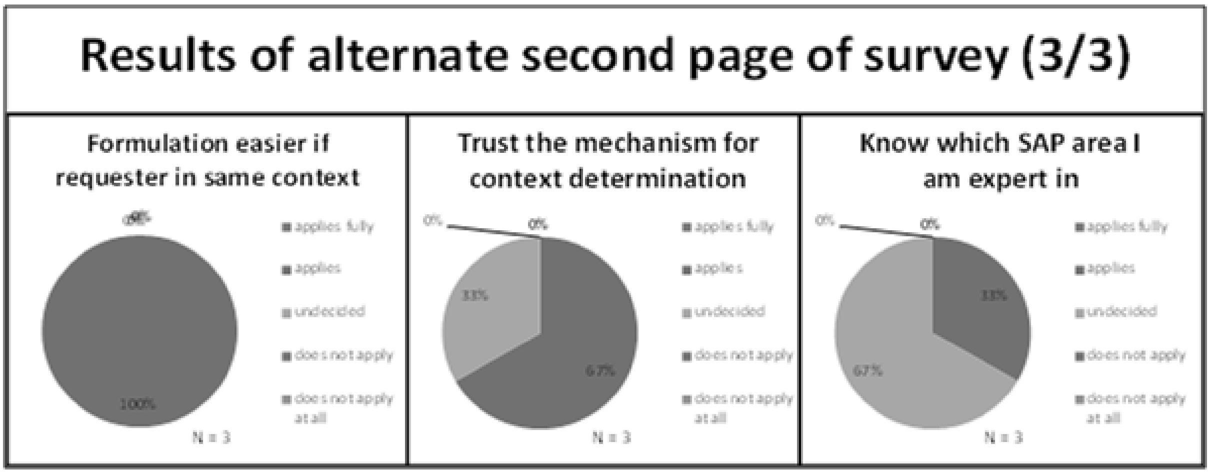


Figure C-21 ***Results of the second case study's survey – alternate second survey page, part 3 of 3***
Source: Own illustration

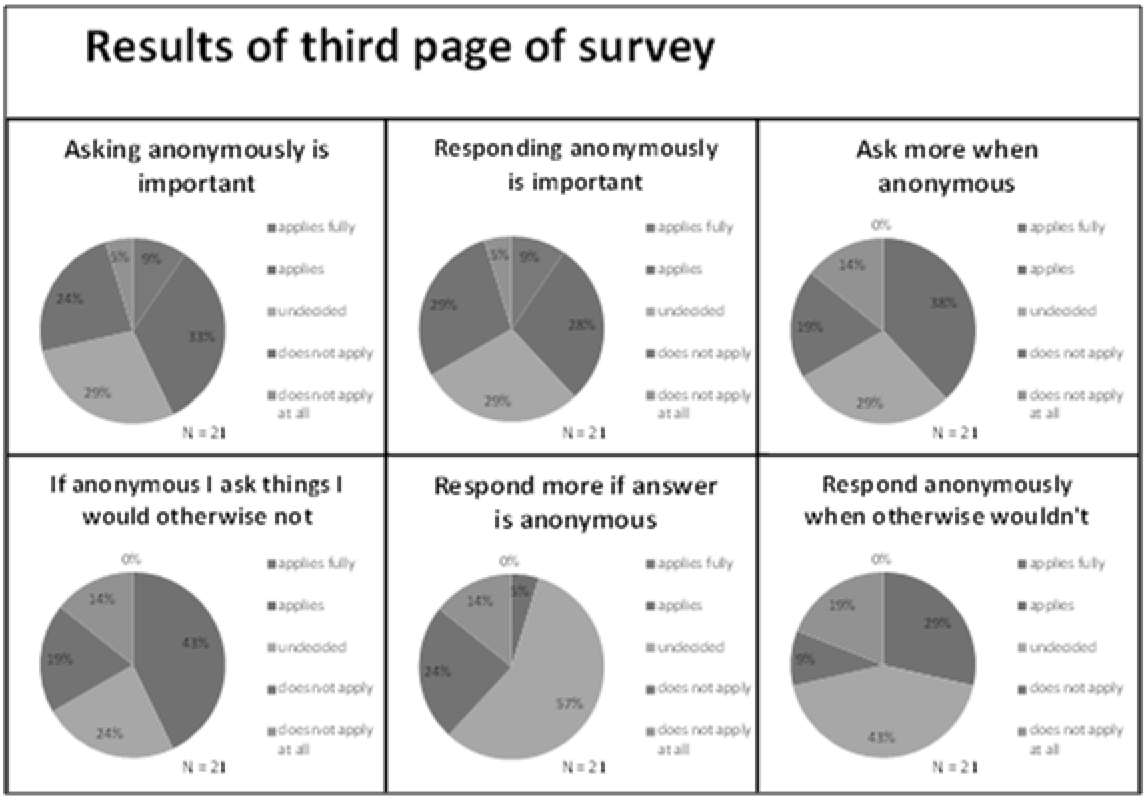


Figure C-22 ***Results of the case study's survey – third survey page***
Source: Own illustration

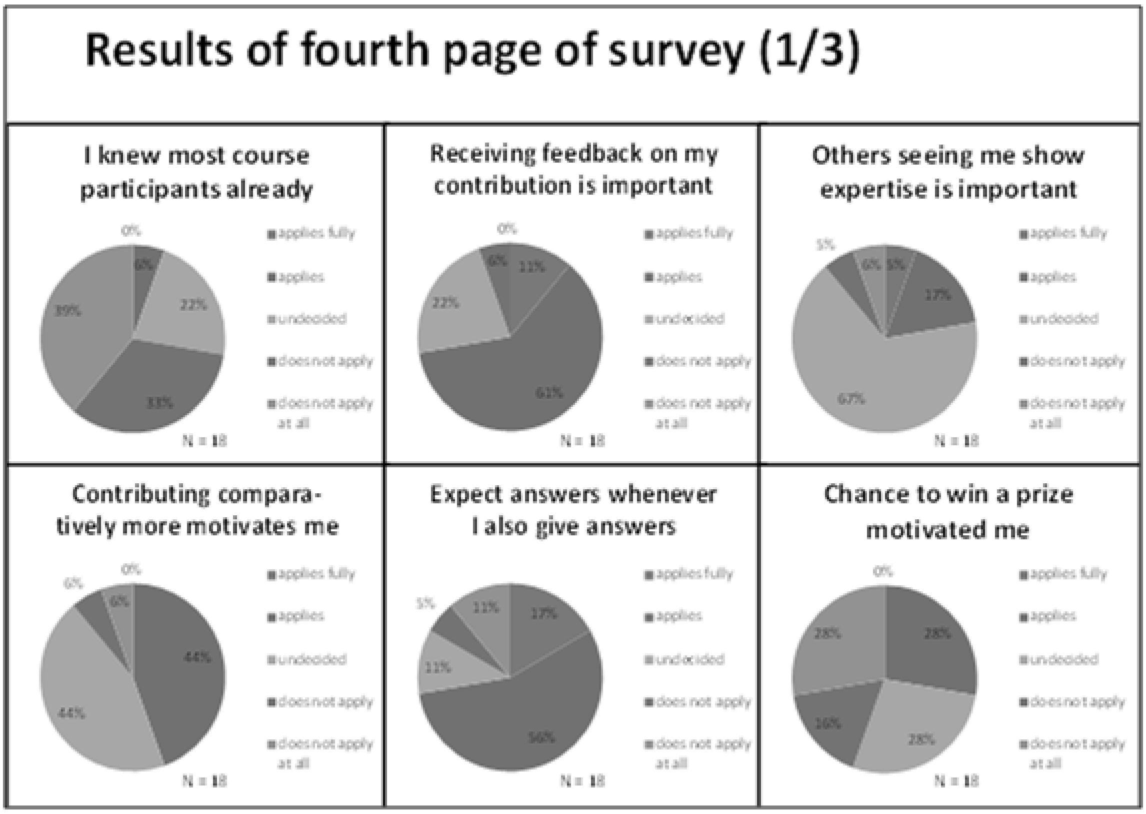


Figure C-23 ***Results of the case study's survey – fourth survey page, part 1 of 3***
Source: Own illustration

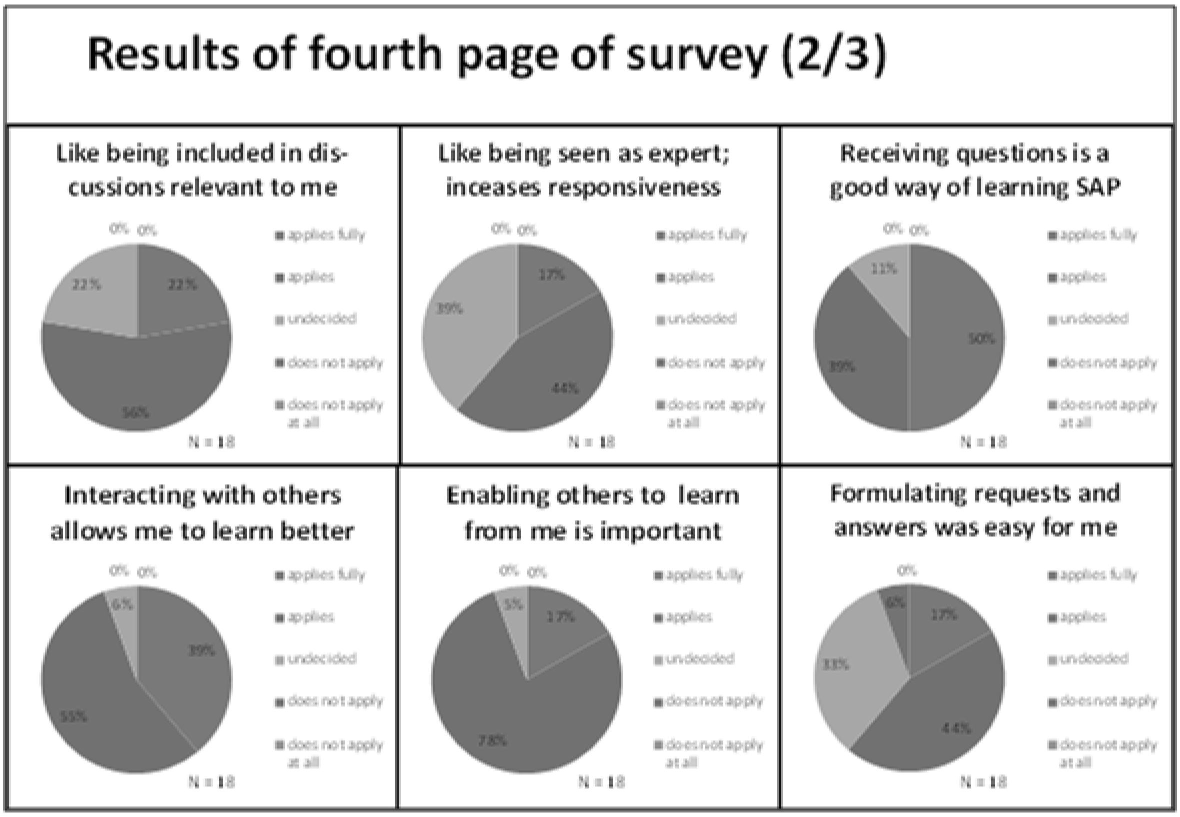


Figure C-24 ***Results of the case study's survey – fourth survey page, part 2 of 3***
Source: Own illustration

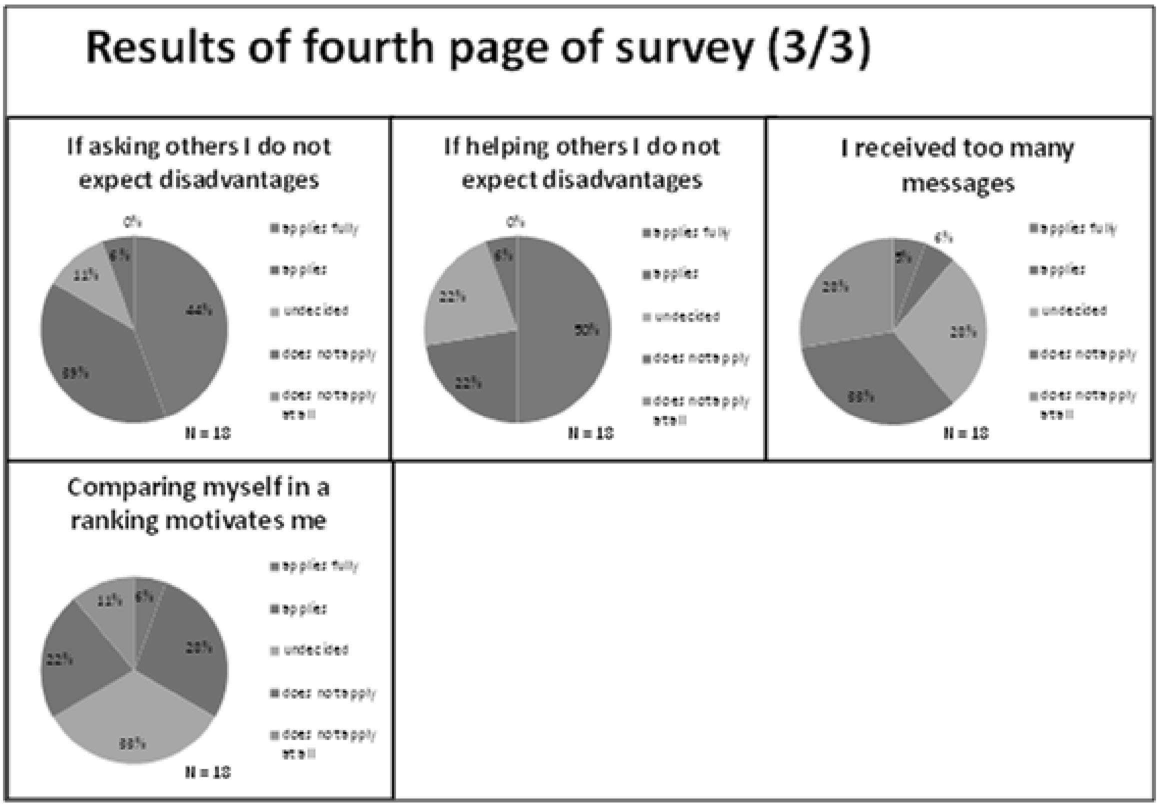


Figure C-25 ***Results of the case study's survey – fourth survey page, part 3 of 3***
Source: Own illustration

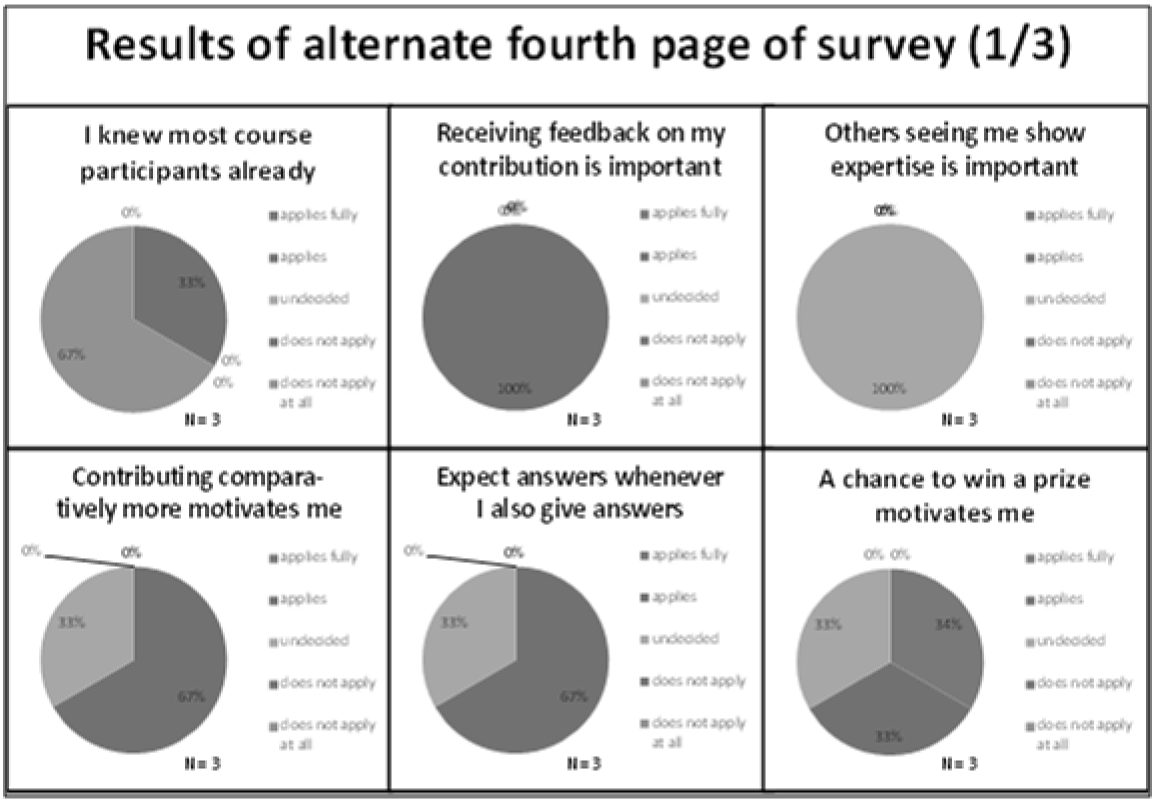


Figure C-26 ***Results of the case study's survey – alternate fourth survey page, part 1 of 3***
Source: Own illustration

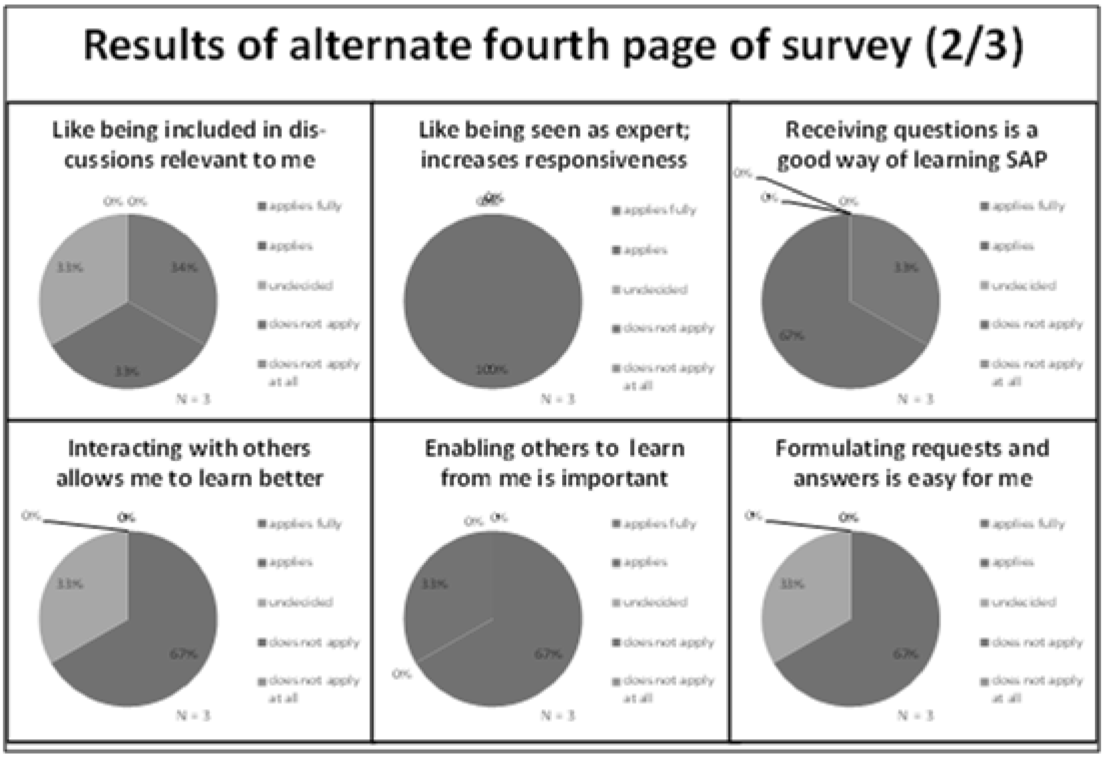


Figure C-27 ***Results of the case study's survey –alternate fourth survey page, part 2 of 3***
Source: Own illustration

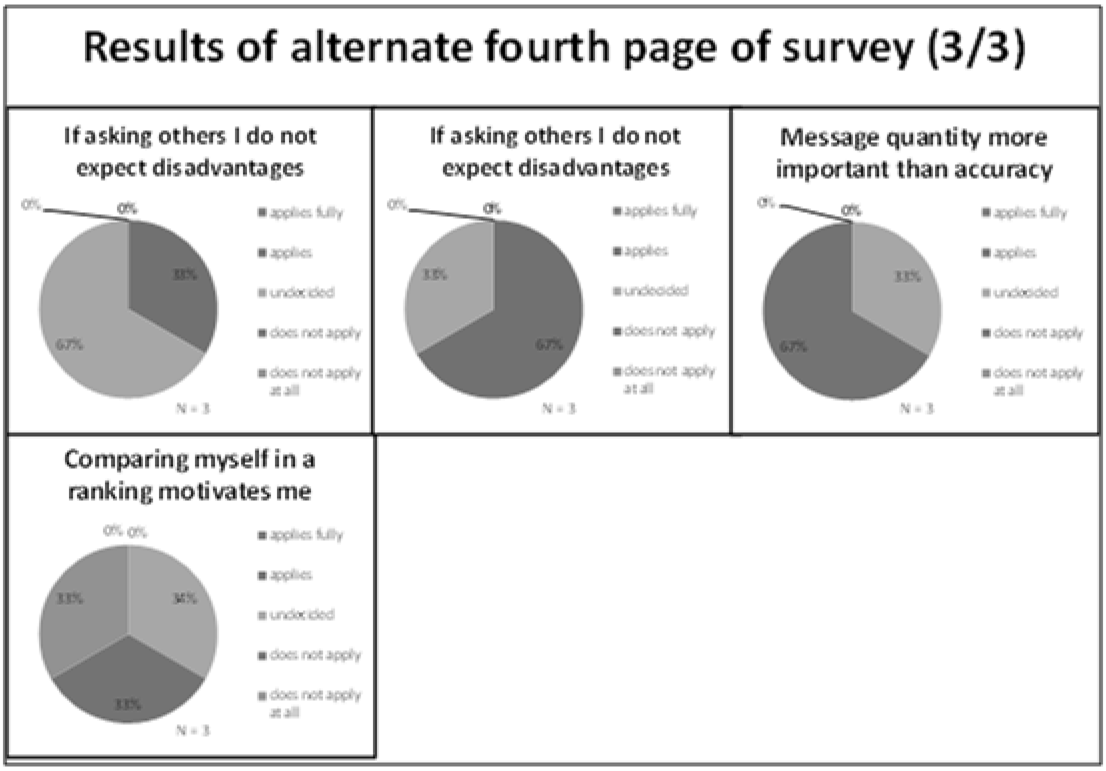


Figure C-28 ***Results of the case study's survey – alternate fourth survey page, part 3 of 3***
Source: Own illustration

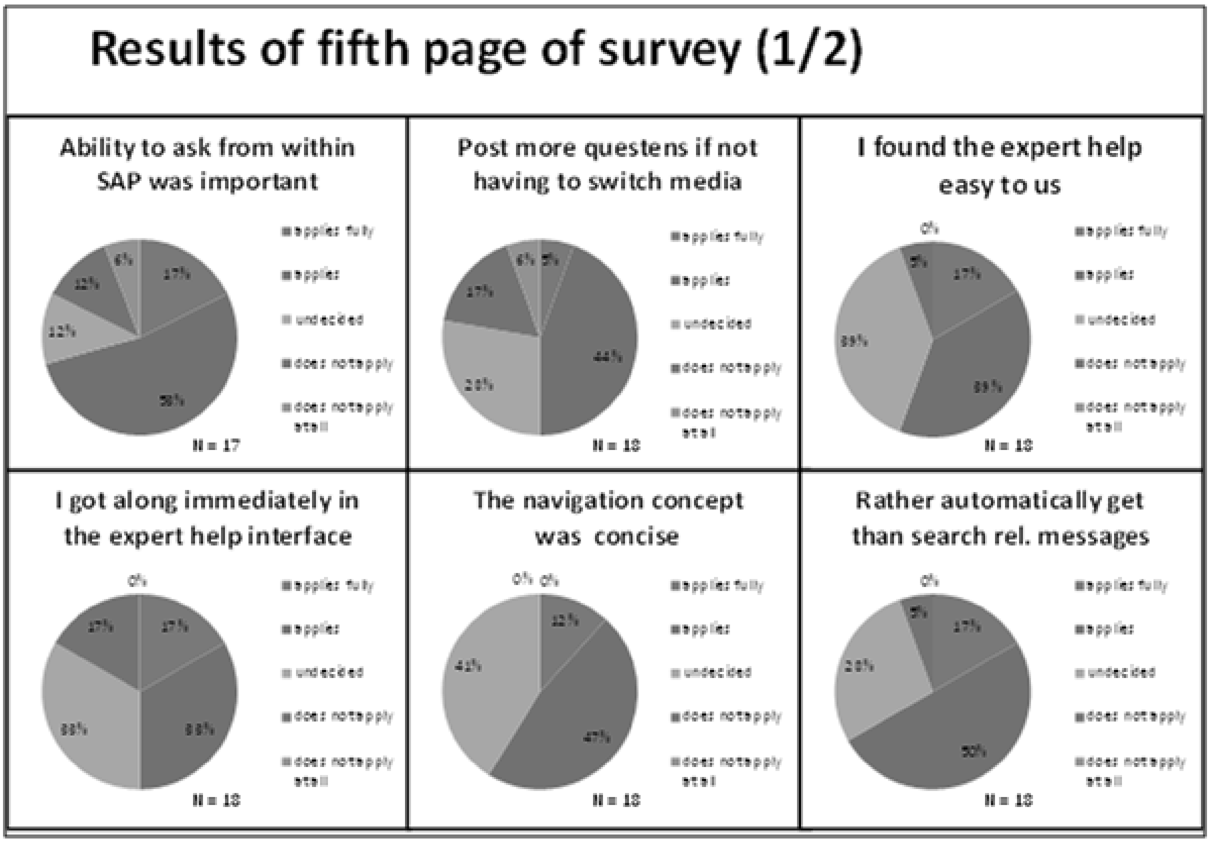


Figure C-29 ***Results of the case study's survey – fifth survey page, part 1 of 2***
Source: Own illustration

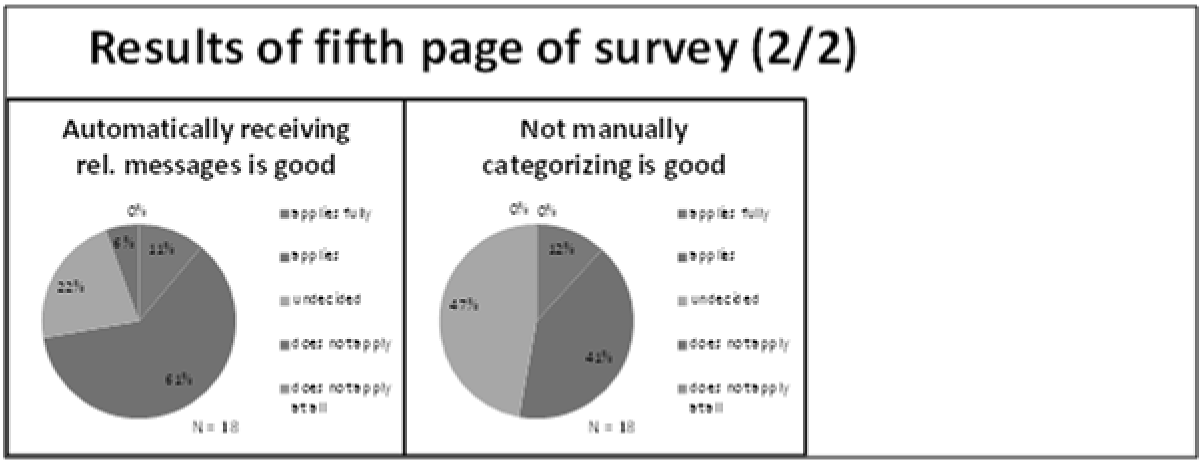


Figure C-30 ***Results of the case study's survey – fifth survey page, part 2 of 2***
Source: Own illustration

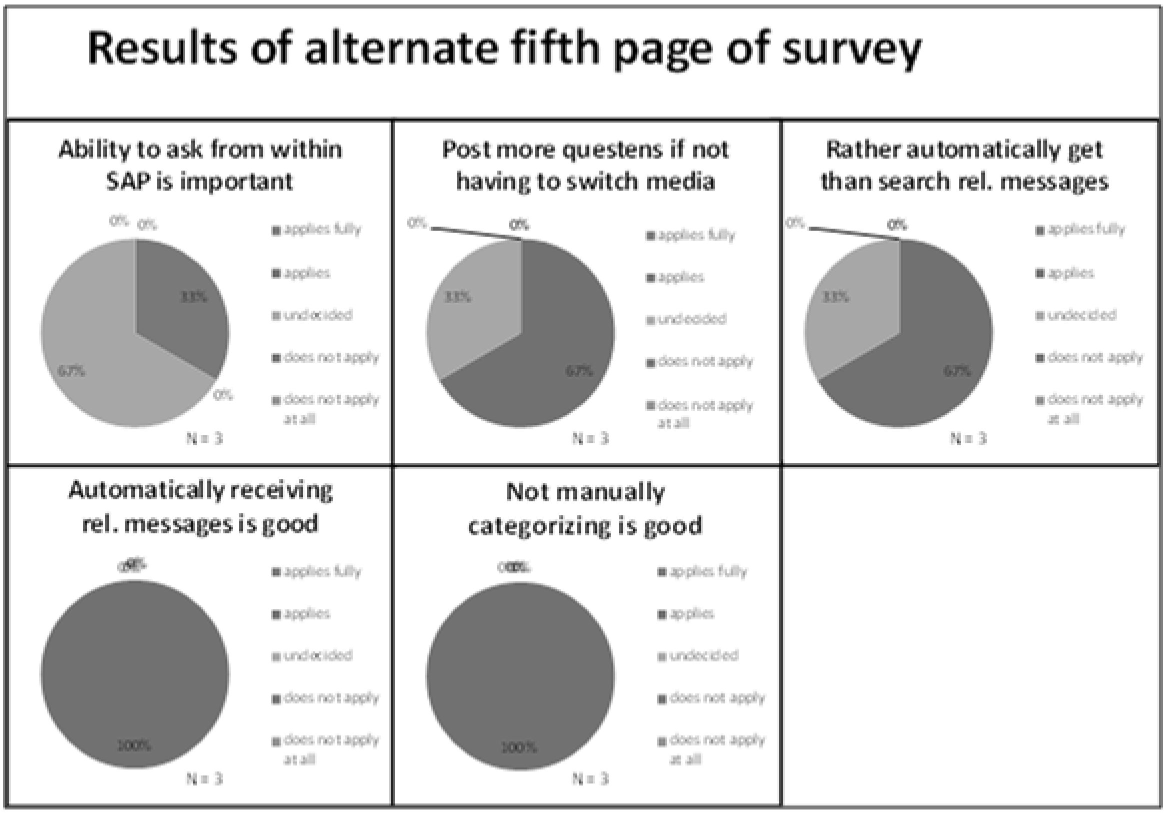


Figure C-31 ***Results of the case study's survey – alternate fifth survey page***
Source: Own illustration

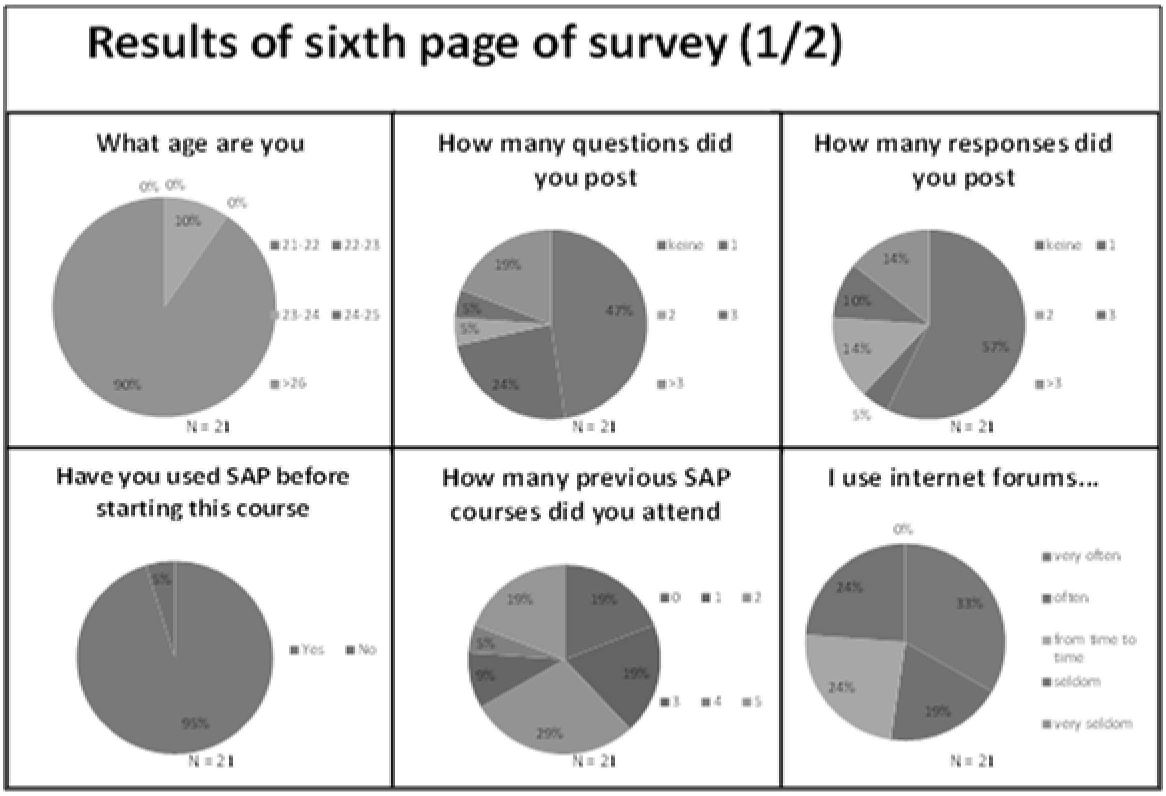


Figure C-32 ***Results of the case study's survey – sixth survey page, part 1 of 2***
Source: Own illustration

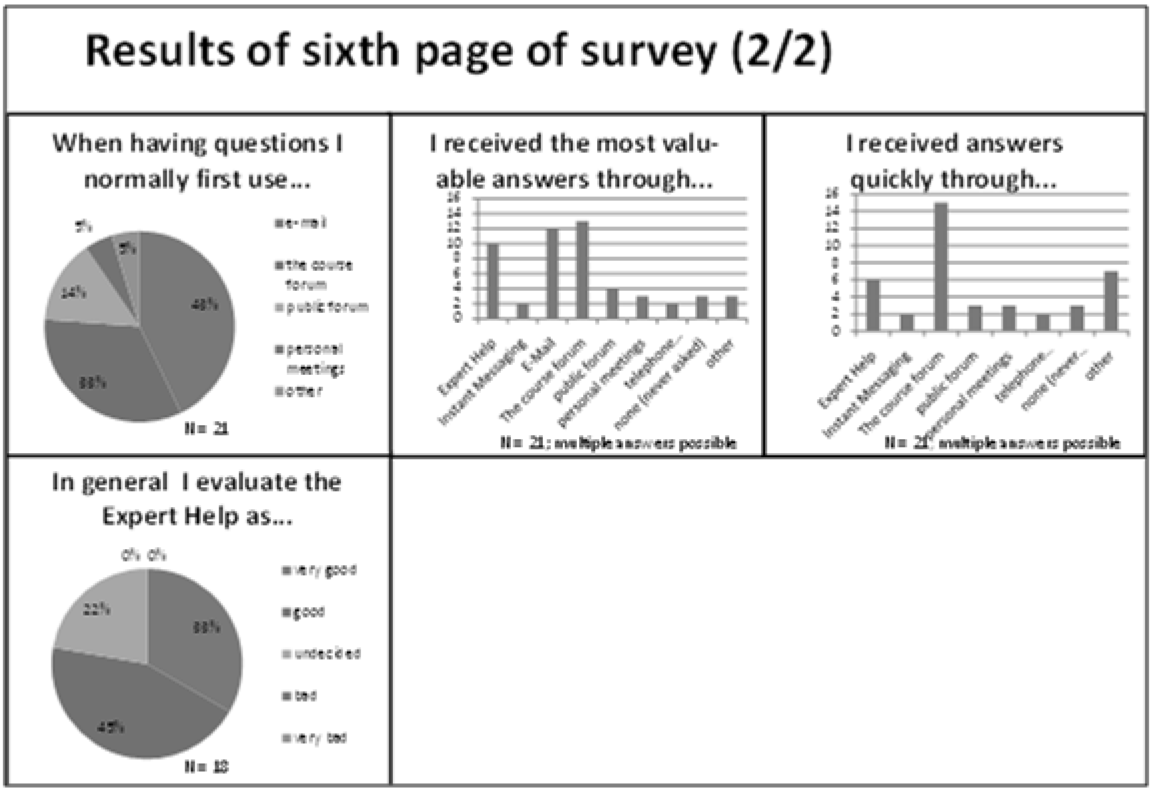


Figure C-33 ***Results of the case study's survey –sixth survey page, part 2 of 2***
Source: Own illustration