Wiki Communities in the Context of Work Processes

Frank Fuchs-Kittowski
Fraunhofer ISST
Mollstraße 1
D-10178 Berlin, Germany
+49 30 24306-326
frank.fuchs-kittowski@isst.fhg.de

André Köhler
University of Leipzig
Klostergasse 3,
D-04109 Leipzig, Germany
+49 341 97-32335
koehler@ebus.informatik.uni-leipzig.de

ABSTRACT
In this article we examine the integration of communities of practice supported by a wiki into work processes. Linear structures are often inappropriate for the execution of knowledge-intensive tasks and work processes. The latter are characterized by non-linear sequences and dynamic social interaction. Communities of practice, however, often lack the „guiding light" needed to structure their work. We discuss the primary requirements for the integration of formally described knowledge-intensive processes into the dynamic social processes of knowledge generation in communities of practice and use the wiki approach for their support. We present our approach for an appropriate interface to integrate wiki communities into process structures and an information retrieval algorithm based on it to connect the process-oriented structures with community-oriented wiki structures. We show the prototypical realization of the concept by a brief example.

Categories and Subject Descriptors
H.5.3 [**Group and Organization Interfaces**]: Collaborative computing, Asynchronous interaction, Computer-supported cooperative work, Web-based interaction. H.3.3 [**Information Search and Retrieval**]: Relevance feedback.

General Terms
Algorithms, Design, Human Factors.

Keywords
Ontology, wiki, community, cooperative knowledge generation, knowledge work, work processes, knowledge processes, process-oriented knowledge structures

1. INTRODUCTION
The knowledge to handle concrete, practical tasks in the work process often already exists in corporate IT systems, e.g. in information systems, documents, and community applications, and should be transparent and available throughout the organization via a collectively accepted and work process oriented knowledge structure. Process structures are particularly suitable for this purpose because it is easier to establish a relationship between the company's own work processes and the available relevant knowledge carriers (e.g. people, documents etc.) [1], making the latter's location and use for processes and other specific activities easier. Technologies and tools for process-oriented knowledge management can provide support for such a structured supply of knowledge; **APO**PILOT is an example [2].

However in knowledge-intensive work processes, it can also be necessary to develop new knowledge ad hoc and in cooperation with other people. Current IT solutions for knowledge management cannot do this, though, because they only make extant, explicit knowledge, e.g. in the form of documents, accessible using inflexible, context-based knowledge structures (ontologies).

Both knowledge-intensive tasks and processes that generate knowledge are nonlinear, dynamic, and socially embedded [3]. In particular, process workflows and the concomitant required knowledge cannot be completely determined and made available in advance. In contrast, communities support communication that is essential for the exchange of knowledge and experience, making processes for cooperative knowledge generation and problem solving possible. Knowledge generated and made explicit in communities, though, is often characterized by its chaotic structure, enormous size, and rapid growth, so any tool aspiring to supply technology-based support - such as a wiki - must meet the challenge of providing context-based access to it.

In this paper we propose a concept for the support of cooperative knowledge generation for knowledge-intensive work processes in wiki communities and a tool based on it. Our goal is the integration of wiki communities with process-oriented knowledge structures. We design a community support mechanism that refers to the work process as a context and focusses on the tasks and problems that can occur. Thus on the one hand, technology supports context-based access to the community's knowledge and on the other it supports the flexible generation and conservation of new knowledge and experience alongside knowledge structures. This will help overcome two weaknesses: the lack of context-based access to the chaotic knowledge generated in today's (wiki) community applications and inadequate flexibility for cooperative knowledge generation in knowledge management systems.

Knowledge networks, technical tools for the electronic storage of the community's knowledge, help store and supply knowledge to community members, who can combine their knowledge with other knowledge network components. Furthermore, members of the community have access to knowledge stored by other members. A wiki is an example of this type of knowledge network.

We discuss the requirements of a software tool that addresses our goal described above in Section 2, and in Section 3 we demonstrate that the wiki approach meets the demands for cooperative knowledge generation in communities. Following in Section 4, we

WikiSym '05, October 16-18, San Diego, CA, U.S.A.

develop our approach for the integration of communities into process-oriented knowledge structures, for which an interface has been conceptualized; it provides process-specific and related context-specific view on the community knowledge. The conception of this interface generates synergies from the connection between communities and the process-oriented knowledge structures. Finally, in Section 5 we present the implementation of this approach in the form of a prototype.

2. Connecting Process Structures to Wiki Communities

Both the provision of knowledge and its cooperative generation are necessary for knowledge-intensive work processes, and to achieve this, knowledge management technologies have to be appropriately linked to community support tools [4]. Furthermore, this integration must insure the continuous cooperative development of the structured knowledge provided.

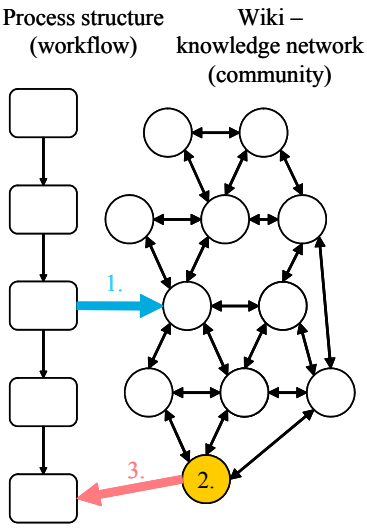

Figure 1: The connection of communities with work processes

This necessitates the creation of a tool that can also meet the following additional requirements (see Figure 1):

1. *The provision of context-related (process-oriented) access to the knowledge network and community knowledge*. The specific situation sets the context for the interpretation of available information. This context in turn can be used for the creation of a process-specific perspective of the knowledge network, establishing the context-specific relevance of the network's individual components.
2. *The process spanning, cooperative generation and conservation of knowledge in the community*. This must occur without the designated process structures restricting the community's social, self organized knowledge generation process. This basically means that an extension of the community knowledge can be achieved in a way that spans processes, making an (explicit) allocation to individual process steps unnecessary.
3. *The possibility to (loosely) associate knowledge components.* The cross-process construction of the knowledge network may lead to a community member's failing to perceive the relevance of network components not directly linked with a process step. This missing reference to the work process (context) can result in a diffuse structure that causes the user to lose orientation in the knowledge network. Thus, the context (i.e. reference to the individual activity in the work process) of any component in the network must be retraceable.

Our approach meets these integration requirements; it is discussed in Section 4.

3. Using the Wiki Approach in Community Support for Cooperative Knowledge Generation

The wiki approach appears to be a suitable solution for IT support of cooperative community knowledge generation. A wiki-wiki web is an open authoring environment for creating and maintaining a community knowledge base, offering a quick, simple way to produce and review information that can be gathered and linked to other wiki pages. All users can comment on, change, supplement, and even delete the wiki's pages. Producing new pages is quick and easy, as is linking them to existing ones.

The wiki's winning feature, however, is that it provides a simple means of interaction due to the simplicity with which users can navigate its pages. The line between the „active" content purveyors (authors) and the „passive" users is largely eliminated, resulting in the rapid appearance of a „chaotic" (knowledge) network of wiki pages and sites.

It is a technology eminently suited for the creation and evolution of knowledge in communities. This knowledge can easily be modified and expanded, permitting the establishment of an integrated, interdisciplinary, and cooperative knowledge base. The central access some or all users have makes the wiki appropriate for project work, document production, the joint development of project concepts, and discussion forums of all types.

Wikis' features also give good reason to apply them in a variety of other scenarios such as content management systems, discussion boards, or other forms of groupware support. With the help of a wiki, users can easily gather and integrate knowledge into the existing (wiki) knowledge base by the user. The particular advantage of the wiki approach compared to other cooperative knowledge generation and exchange systems is the focus on the *process* as well as the *result* of communication (see Figure 2).

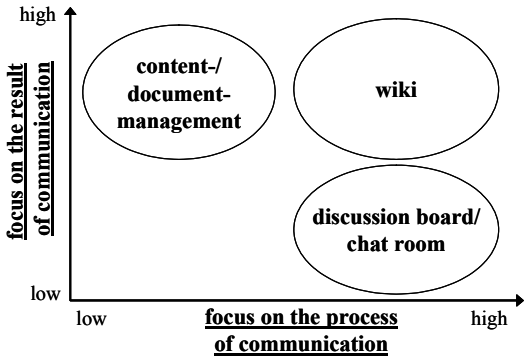

Figure 2: Communication-oriented classification

Content and document management systems tend to focus on the exchange of results of tasks done by several people. Communication processes and discussions that lead to the „end product", the

document, are limited, for example, to annotations. Discussion boards, in contrast, focus predominantly on the cooperation process making the exchange of opinions and a general consensus regarding the document possible. The results of the discussion are generally implicit in the individual postings and have to be extracted and condensed afterwards. Wikis, in contrast, permit users to discuss *and* work on the result simultaneously.

The wiki's features largely fulfill the afore-mentioned requirements for a tool to support cooperative knowledge generation. Moreover, cooperative production of content becomes very efficient through the realignment of the distinction between author and reader.

Wikis also contribute to participants consciously and actively benefiting from both the knowledge of others and the creation of group consciousness that are crucial in the community building process. This is due to the feeling of working jointly with other experts forged by the member's perception of her or his own participation, as well as through the absence of a supervisory monitoring authority.

4. The Integration of Knowledge Networks into Process-oriented Structures

In the following we present the concept underlying the integration of knowledge networks into process-oriented knowledge structures. For this purpose we design an interface providing a process specific and thus context specific view of the knowledge network. The approach is based on the idea of including the process structure in the development of the knowledge network as one of the design criteria.

The assumption that users will have a sense of being in a familiar process step when knowledge gaps occur forms an important basis for the provision of information extracted from the network. Incorporation of the users' particular situation creates the context in which they interpret the supplied information. This context can in turn be used to give a process-specific perspective of the knowledge network, allowing a context-specific assessment of the network's components. This supports the community in the process of cooperative problem solving and knowledge creation and furthermore supports the documentation of the solution, i.e. the result of the process.

This integration concept results in a formal description of the interaction between a process model, a knowledge network (wiki), and a corresponding interface.

4.1 Modeling an Interface between a Process-oriented Structure and a Knowledge Network

In order to foster a process-specific perspective of the knowledge network, it is necessary to anchor a component that is solely associated with the corresponding process (step) in the network. This component creates a one-to-one connection between a specific process and any selected network component. Figure 3 explains the structure of this type of interface.

In the lower part of the diagram, processes P1 and P2 are depicted as a section of a knowledge-intensive process. Both have been sub-divided into two sub-steps (activities), T1 and T2. The arrows connecting the processes or the sub-steps show the sequence of their execution. Figure 3 illustrates an example with two levels of hierarchy. In principle, any number of levels is possible.

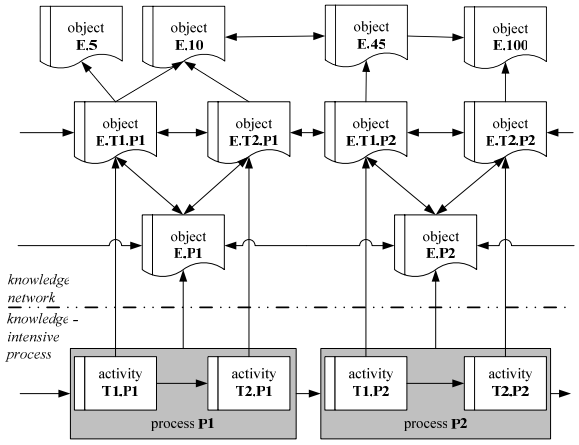

Figure 3: Shaping an interface between the process model and knowledge network

The arrows connecting the processes (or sub-steps) with the components of the knowledge network represent the interface itself. Every part of the knowledge-intensive processes can be directly ascribed to a network component. By means of these connections, the process structure in the network can be automatically reproduced if necessary. In addition to the referential links to illustrate the process structure, the network components primarily contain links to those knowledge components of the network relevant for the individual processes. The development of this type of knowledge components and the creation of corresponding links are subject to modification by network users.

4.2 Information Retrieval in the Integrated Knowledge Network

The interface linking knowledge-intensive processes and the knowledge network permits users to adjust the information retrieval in accordance with their information requirements and to evaluate the network components correspondingly.

There are two basic assumptions here. The first is that two documents referring to each other have a contextual correlation [5], [6], and the second is that an inverse relationship exists between a document's relevance for a specific process step and the number of references between the process step and the document. It is recognized, by the way, that these assumptions describe an ideal type of network conditions and might not apply in real life circumstances.

For a better understanding of this situation, the term *referential distance* is introduced here. In the following, *D* is the number of available documents and *P* the number of process steps. If the relevance of a document *d* to a process step *p* is r_{pd} for p = 1...P and d = 1...D, the *referential distance* t_{pd} represents the distance of the document to the origin of the reference chain, the process *p*. The referential distance corresponds to the shortest distance between the point of origin and the correlating document measured by the number of the *minimally* necessary references. In this way it is possible to describe the potential relevance of a document compared to the origin of referencing. If the referential distance increases, the relevance can be expected to decrease.

If the entire knowledge network is transformed from a process-specific perspective, the referential distance t_{pd} compared to the current process step or the interface component must be deter-

mined for every document *d* from the outset. Subsequently, all the documents with the same reference distance t_{pd} can be pooled in what we term *distance classes* k.

Regarding an extensive knowledge network with thousands of pages, it is easy to see that the referential distance can only roughly structure the relevance of documents. Since documents in the same distance class may differ considerably, a more precise distinction becomes inevitable. We assume that the relevance of a document in a distance class increases with the number of references from the same or a higher-ordered distance class. This is based on the assumption that the reference to a document can be interpreted as a vote for this document.[1] A rating of a document in this sense is conducted solely on the basis of a vote of other documents with at least equal or higher relevance (based on the referential distance). This term is be defined as *distance class-depending frequency of reference* h_{kd} within a distance class k.

These specific distance classes k and the corresponding frequencies of reference h_{kd} serve as a means for the process-specific sequencing of documents. The following function is applied in order to determine the relevance r_{pd}:

$$r_{pd}(k, h_{kd}) = \begin{cases} \dfrac{1}{k - \frac{h_{kd}}{m}} & \text{if} \quad k - \dfrac{h_{kd}}{m} \geq 1 \\ 1 & \text{if} \quad k - \dfrac{h_{kd}}{m} < 1 \end{cases}$$

Formula 1: Relevance

Because of their inverse relationship, the value of the relevance initially decreases as distance class *k* increases. It can, though, be improved by references h_{kd} to the document. *K* is therefore reduced by the quotient $\frac{h_{kd}}{m}$. The parameter *m* indicates the number of references that would help to increase the relevance of a document exceed its original distance class. The relevance of a document d = 1 from the distance class *k* with h_{k1} = m would correspond to the relevance of a document d = 2 from the distance class k-1 and $h_{(k-1)1} = 1$. If $h_{k1} > m$, the relevance of the document *d* = 1 has a value corresponding to the relevance of a document *d* = 2 with $h_{(k-1)2} < m$. The relevance of a given document can have a value ranging from 0 to 1, with $r_{pd} = 1$ being the highest possible relevance. If the number of references is increasing so that $r_{pd} > 1$, $r_{pd} = 1$ is assumed. The interface document always has the relevance $r_{pdS} = 1$.

The search algorithm is illustrated in the following. We assume users to be in the process of „developing a basic concept (Basic Concept)“ and searching the wiki for the term „requirements specifications“ as part of the process. Figure 4 lists all the processes and wiki pages relevant for this process.

Figure 4 also shows where in which distance class *k* the individual pages are located from the point of view of the process „developing a basic concept“. Moreover, the frequency of reference h_{kd} as well as the process-specific relevance r_{pd} of every page is revealed, based on the parameter *m* = 3. All the documents shown in Figure 4 (black and grey) are searched for this term.

[1] This assumption is partly founded on the PageRank™-Technology by Larry Page and Sergey Brin, the core piece of the search engine Google [19].

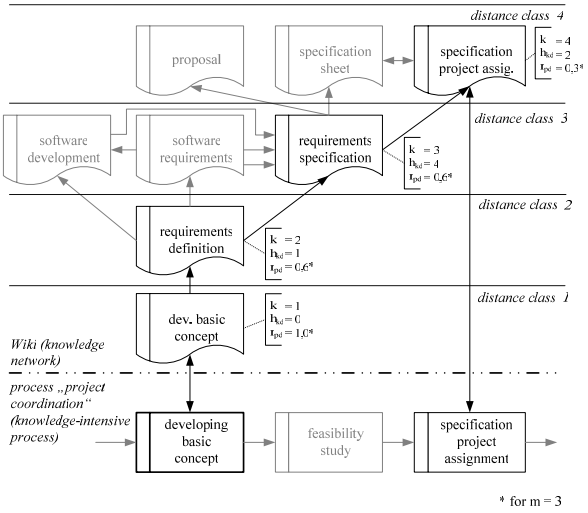

Figure 4: Diagrammatic illustration of a section of a wiki for the process „developing a basic concept“ (explicit query)

Based on the distance class k = 1, calculating the relevance of the document „developing a basic concept“ yields the result $r_{pd} = 1$. The frequency of reference h_{kd} is zero since this page is not referred to by any other wiki page and reference from the process is only made in an imaginary way because of the matching process described above. The document „assessment analysis” is located in the distance class k = 2 and contains only one link resulting in a relevance $r_{pd} = 0.6$. The document „requirements specifications“ located in the distance class k = 3 has the same process-specific relevance, $r_{pd} = 0.6$. This value, which is untypical for the distance class, occurred because of the number of references $h_{kd} = 4$. The document „surpasses“ its distance class since, in this case, $h_{kd} > m$ and receives the same relevance as the document „assessment analysis“, which is of a higher order in terms of process. The document „specifying the project instructions“ is marked with k = 4 and $h_{kd} = 2$ resulting in a relevance $r_{pd} = 0.3$. This result will be graded and presented in accordance with the relevance, as can be seen in Figure 8.

5. The Realization of the Prototype

In this section we will introduce a tool that meets the specifications described above. This implementation consists of three independent components. First, **APO**PILOT is used as a process-oriented, organisational knowledge base. Second, the PmWiki will serve as the Community Support System for the process of cooperative knowledge generation. Furthermore, in accordance with the above-mentioned concept, an interface between **APO**PILOT and PmWiki has been implemented.

In the following, **APO**PILOT and PmWiki are briefly introduced. Subsequently, the interface will be presented in detail.

5.1 Process-oriented Knowledge Base – **APO**PILOT

The **APO**PILOT [2], a tool that accompanies work processes, consistently applies a process orientation approach. The **APO**PILOT permits process-oriented navigation through a firm’s work processes that have been modeled. As an assistant without an active control component, it visually presents the process flow in the form of event-process chains and provides descriptions of

processes, sub-processes and process steps to underpin the workflow, aiding the structuring of the work process.

To every process step and activity **APO**PILOT provides corresponding sources of knowledge to help employees cope with their tasks: In the „Library" domain, documents and other relevant learning materials (e.g. from the corporate Intranet) are provided. The „People" domain identifies and suggests individuals qualified to act as competence carriers and includes means of communicating with them: email, telephone, videoconference, etc. There is also a „Discussion" domain with moderated forums for exchanging experience, points of view, and opinions, as well as trouble-shooting tips.

Access to a wiki for cross-process support of cooperative knowledge generation is also available.

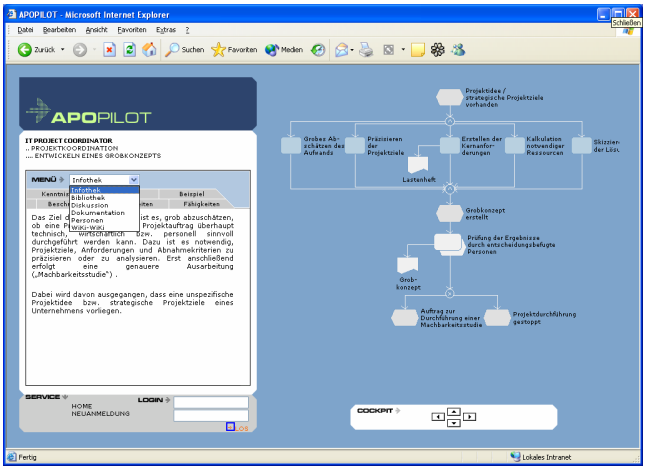

Figure 5: APOPILOT

These sources of knowledge, which are usually dispersed throughout various IT systems and independent from work processes, are now integrated and structured in a uniform, process-oriented view.

5.2 Co-operative Knowledge Generation – PmWiki

Patrick R. Michaud's PmWiki [7], a small wiki developed in PHP, was adapted for use in **APO**PILOT. It was chosen because of its availability under a GNU Public License Agreement and its simple adaptability due to its being written in PHP.

Moreover, PmWiki has a number of features not usually found in wikis that are quite useful in combination with **APO**PILOT. Among these is the possibility to pool documents and provide them with access rights. Another consideration was based on the seemingly well-founded assumption that the application will undergo continuous development, creating stable grounds for the evolution of the prototype introduced here.

5.3 Functionality of the Interface

As a sample of a process-structured knowledge base, **APO**PILOT provides access to the wiki for every process-related step and activity. A corresponding wiki page can be addressed immediately from any activity the user happens to be working on; here „developing a basic concept" is given as an example. Furthermore, the entire wiki can be searched for a specific term (e.g. „requirements specification", see Figure 6).

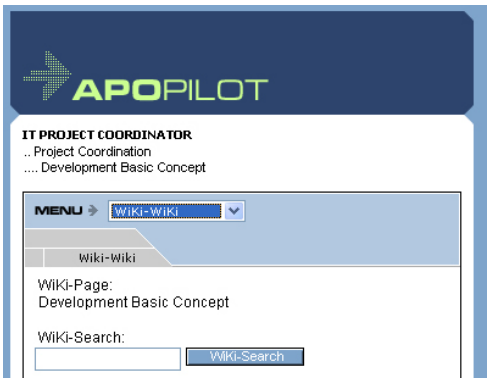

Figure 6: Access page **of the Wiki**

A direct link to the wiki opens the correlating process-specific interface document in a separate browser window. This wiki page can be edited, new pages can be created, and it is also possible to „navigate" through the entire wiki (see Figure 7).

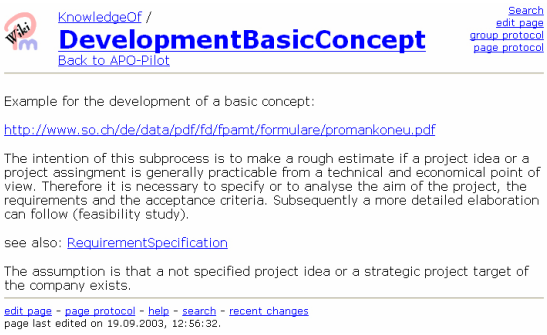

Figure 7: Interface page

A query for a specific term results in a search of all the wiki's pages. With the help of distance classes and the distance-class-dependent frequency of reference of the process step selected, the hits are screened, evaluated, and presented according to their process-specific relevance (see Figure 8).

Besides the relevance, expressed graphically by a green bar and also given as a percentage, the result contains features of the correlating distance class („distance"). The name of the page and the context of the search key are presented as well. Clicking its title opens the wiki page, and at the top, general information about the query result is given, particularly regarding the process that was its subject. This is important since a query with the same search key for another process would lead to a completely different result. In addition, a tool that explains the query result can be called.

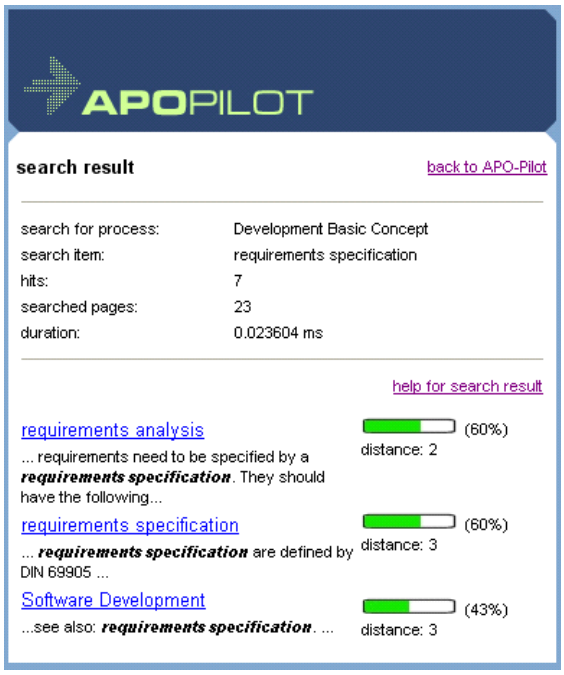

Figure 8: Process-related result

It is possible to return to **APO**PILOT from every wiki page. If the current page is part of the interface, the corresponding process step is opened in **APO**PILOT. If no matching page is found, all the process steps with the shortest referential distance from the page selected are located and as a link to return to **APO**PILOT provided (see Figure 9). A link to return to the process step started from in **APO**PILOT is also available.

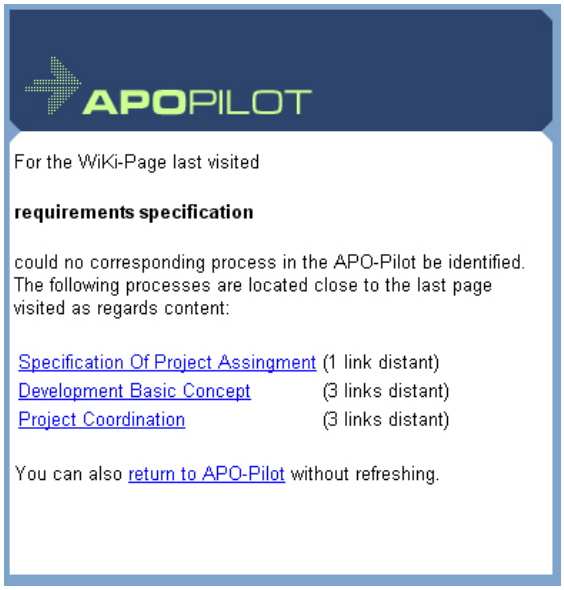

Figure 9: Return to the process model

5.4 Technological Design of the Interface

Linking **APO**PILOT with PmWiki has been done using HTTP based communication. To do this, both programs send their request to an interface file written in PHP. A mapping file that contains process identification and the names of corresponding HTML pages from **APO**PILOT is also necessary for the interface file to be able to locate the corresponding page in **APO**PILOT or the wiki based on the process title, or to perform a process-specific query. The interface documents are generated automatically by calling a page with the corresponding process title in the wiki. If the page does not exist, then the page is automatically generated and can subsequently be filled with content. The above-mentioned components can be filed in a distributed way thus achieving the highest possible degree of flexibility. The correlating interface can be adapted accordingly via parameters.

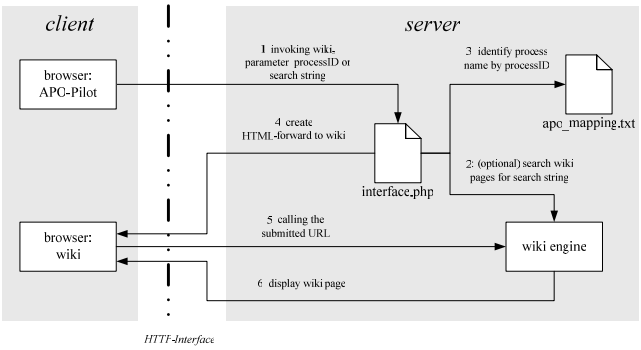

Figure 10: Interaction between the components upon calling the wiki from APOPILOT

In the following we describe the interaction between the components using the example of „Process-related opening of the wiki via the wiki start page in **APO**PILOT“. To start the wiki in a specific process, clicking the wiki link in **APO**PILOT opens a new browser window in which the given URL of the interface file is called. A one-to-one process identification - and search key if needed - is transmitted via parameters. With the help of the mapping file, the actual process title is determined in order to form the correlating wiki term. As a result, the interface file sends an HTML file that includes a forwarding link to the wiki page that has been located to the open browser window. Figure 10 elucidates the interaction between the components.

6. Summary and Conclusion

Knowledge-intensive processes are primarily characterized by problems that cannot be anticipated, which, in addition to a process-related supply of available knowledge sources, also comprise processes for cooperative problem solving and knowledge generation in communities. Due to the highly complex nature of the problems involved, however, cooperative problem solving requires a cross-process exchange and generation of knowledge. We presented the wiki approach as a technological solution for the creation of a community to address this task.

We described an integration concept for the connection of a wiki community and a process-oriented knowledge structure. Our goal is to create synergies from the two approaches. The core piece of our approach is an interface that makes both the generation of process-specific perspectives of community knowledge (knowledge network) on the basis of the referential distance possible as well as identification of the document's distance-class-dependent frequency of reference. Last but not least, we also presented a prototype implementation designed on the basis of the integration approach. We integrated wiki into the application **APO**PILOT to achieve this.

In the following we outline some other aspects that might be considered for inclusion in the further development of the approach that has been presented:

6.1 Community-dependent Wikis

It might be worth examining whether different wikis should be created for different groups of users. This approach might be interesting for maintaining knowledge dissemination in large communities. In exceedingly large communities, the individual may not feel recognized. There could also be difficulties in mutual understanding due to a wiki's having too many domains of specialization. Individuals being perceived as experts, though, is one of the central requirements of a working Community of Practice [8], [9].

6.2 Standardized Links

The associative links used in wikis signal semantic similarity between two documents. Unfortunately, the nature of the documents referred to frequently remains unclear: it cannot be ascertained whether the document contains a definition, a discussion of the topic, a free contribution, etc. Using standardized links, the user can be informed about the type of document that can be expected *before* clicking the link. This could be realized by introducing a number of easy-to-understand types of documents (definition, discussion, etc.). Corresponding possibilities are offered by the HTML 4.0 standard [10].

6.3 Using Topic Maps to Structure the Wiki

Topic maps can be utilized to classify documents in large wikis. The resulting meta-level can be compared with a community-dependent glossary. Wiki users would have to maintain this glossary and in the process systematize the community-specific terminology. The topic map could then contain references to the corresponding wiki pages (occurrences) [11].

7. REFERENCES

[1] Fuchs-Kittowski, F.: Kooperative Wissenserzeugung und -nutzung in wissensintensiven Geschäftsprozessen. In: Stumme, G.; Schnurr, H.-P.; Staab, S.; Studer, R.; Sure, Y. (Eds.): Professionelles Wissensmanagement - Erfahrungen und Visionen. Aachen: Shaker Verlag, 2001, p. 125-130.

[2] Fuchs-Kittowski, F.; Manski, K.; Faust, D.; Prehn, M.; Schwenzien, I.: Arbeitsprozessorientiertes E-Learning mit Methoden und Werkzeugen des prozessorientierten Wissensmanagement. In: Bode, A.; Desel, J.; Rathmayer, S.; Wessner, M. (Eds.): Proceedings of 1st "Deutsche e-Learning Fachtagung Informatik (DeLFI03)". Lecture Notes in Informatics (LNI), Volume P-37, Bonn: Gesellschaft für Informatik, 2003, p. 392-401.

[3] Brown, J. S.; Collins, A.; Duguid, P.: Situated Cognition and the Culture of Learning. In: Education Researcher 18, (1989), 1, p. 32-42.

[4] Fuchs-Kittowski, F.: Wissensmanagement und E-Collaboration. In: Proceedings of KnowTech 2003. 5th Conference on Knowledge Engineering & Management, Munich, 20.-21. October 2003.

[5] Chakrabarti, S., Srivastava, M., Tiwari, M.: Using Memex to archive and mine community Web browsing experience. 9th International World Wide Web Conference, 2000. In: http://www9.org/w9cdrom/98/98.html.

[6] Croft, W., Turtle, H.: A retrieval model incorporating hypertext links. In: Proceedings of the ACM Hypertext'89, 1989, p. 213–224.

[7] http://www.pmichaud.com/wiki/Experimental.

[8] Lave, J., Wenger, E.: Situated Learning: Legitimate Peripheral Participation. Cambrigde: University Press 1991.

[9] Maslow, A.H.: Motivation and personality. New York: Harper & Row, 1970.

[10] HTML 4.0 Specification, W3C Recommendation, 1998, http://www.w3.org/TR/1998/REC-html40-19980424/.

[11] Pepper, S.: The TAO of Topic Maps: finding the way in the age of infoglut. XML Europe 2000. In: http://www.gca.org/papers/xmleurope2000/pdf/s11-01.pdf.

[12] Brin, S., Page, L.: The Anatomy of a Large-Scale Hypertextual Web Search Engine. Computer Science Department, Stanford University, 1998. In: http://www-db.stanford.edu/~backrub/google.htm