

Association for Information Systems AIS Electronic Library (AISeL)

ECIS 2006 Proceedings

European Conference on Information Systems
(ECIS)

2006

Neither essence nor accident: Situated knowledge and its importance for the community broker

Bernhard Nett

Fraunhofer Institute for Applied Information Technology, bernhard.nett@fit.fraunhofer.de

Tobias Dyrks

tobias.dyrks@fit.fraunhofer.de

Claudia Mueller

claudia.mueller@fit.fraunhofer.de

Marco Durissini

marco.durissini@fit.fraunhofer.de

Follow this and additional works at: <http://aisel.aisnet.org/ecis2006>

Recommended Citation

Nett, Bernhard; Dyrks, Tobias; Mueller, Claudia; and Durissini, Marco, "Neither essence nor accident: Situated knowledge and its importance for the community broker" (2006). *ECIS 2006 Proceedings*. 98.

<http://aisel.aisnet.org/ecis2006/98>

This material is brought to you by the European Conference on Information Systems (ECIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2006 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

NEITHER ESSENCE NOR ACCIDENT: SITUATED KNOWLEDGE AND ITS IMPORTANCE FOR THE COMMUNITY BROKER

Nett, Bernhard, Fraunhofer Institute for Applied Information Technology, Schloss
Birlinghoven, 53754 Sankt Augustin, Germany, bernhard.nett@fit.fraunhofer.de

Dyrks, Tobias, Fraunhofer Institute for Applied Information Technology, Schloss
Birlinghoven, 53754 Sankt Augustin, Germany, tobias.dyrks@fit.fraunhofer.de

Mueller, Claudia, Fraunhofer Institute for Applied Information Technology, Schloss
Birlinghoven, 53754 Sankt Augustin, Germany, claudia.mueller@fit.fraunhofer.de

Durissini, Marco, Fraunhofer Institute for Applied Information Technology, Schloss
Birlinghoven, 53754 Sankt Augustin, Germany, marco.durissini@fit.fraunhofer.de

Abstract

The paper describes the design of the "Community Broker" as a case study for the development of a tool attempting to support individual web-search strategies by using the community-based nature of a great deal of information exchange and learning. Community experiences affect the understanding, validation and organization of individual knowledge. Communities affecting the representation of knowledge frame the every-day decisions individuals have to take in spite of incomplete, insecure knowledge and the vagueness of communicating in natural languages. As for the often distributed, fragmented and heterogeneous nature of knowledge and related problems of information search in the internet - some of which are described in the paper - technical tools can support information search, when the influence of community-based identities upon knowledge representation and -identification is used as a source of additional information instead of being ignored.

Keywords: communities, situated knowledge, distributed knowledge, expertise, reputation, search tools

1 INTRODUCTION

In the recent time, communities have experienced enormous attention, among others, in knowledge management. The detection of Communities of Practice (Lave & Wenger 1991) has motivated discussion on their possible role in online communities in particular, and on the role of technology in relation to organizational development in general. The community conception underlying the understanding of online communities in informatics is based on IT-mediated interaction, whereas the Communities of Practice concept focuses on a shared identity, repertoire, and enterprise among participants. This tension between the different conceptions can be a basis for conceptual work and empirical analyses – but for misjudgements, i.e., equating forum interaction with Communities of Practice, too.

In this paper, we describe the conception of the Community Broker, a software tool. The Community Broker helps to find relations between websites and the logs of online communities through an automated text matching. In contrast to conventional search engines, the Community Broker requires no submission of keywords by the user. The Community Broker works as a suggestion engine automatically proposing threads of online communities, which could be relevant for a user reading the text on a website. The basis of the relevance assumption is the similarity of the vocabulary between the text on the website and the one in the threads. This seems to be a promising application, if it helps to support human search-strategies for the Internet.

At first glance, the individual perspective on human search strategies has nothing to do with the Communities of Practice concept. However, the idea of the Community Broker is based upon the view that there are not only Communities of Practice in society and that these can be identified by not actively participating observers, but that such identification of community backgrounds is a basic instrument of orientation for the individual, in particular, in certain search activities.

The Community Broker concept, such as other technology, needs conceptual reflection and empirical research. As its life cycle has merely come to the state of a prototype implementation yet there are still no empirical findings to be presented regarding the usage of the Community Broker – this is the reason to present this paper as a work-in-progress. However, the conceptual work related to the Community Broker concept may not only be of interest for a scientific audience, interested in communities and technologies, but may in itself enlarge the opportunities of studies on their relation.

The paper starts with some explanations of the work context, in which the Community Broker has been developed, the VSEK project (Virtual Software-Engineering Competence Centre). It shows the ethnographic research done within the field of small and medium-sized enterprises of the German software branch. This research helped to detect the problems of a target group, developers in small and medium-sized enterprises of the German software branch. Nevertheless, there had to be a counter view from the perspective of technical opportunities in order to define a promising product idea.

In the following, the technical realisation of the Community Broker is described through use cases. Some general reflection on the role of communities for the distribution of knowledge in general, and the Community Broker in particular form the end of the paper.

2 KNOWLEDGE PROCESSES IN ENTERPRISES OF THE SOFTWARE BRANCH

The Community Broker has been developed within the BMBF- (German Ministry of Education and Science) promoted VSEK project, which was to develop textual "knowledge modules" to be presented on a portal (www.software-kompetenz.de) for Software Engineering, and to promote Software Engineering in general, first of all, among small and medium-sized enterprises of the German software branch. In this context, a field study on "knowledge processes in the software branch" (Nett & Wulf

2005) was conducted in spring 2003 by the Fraunhofer Institute for Applied Information Technology (Fh FIT) with software developers of small and medium-sized enterprises in the German software branch. In particular, the interviews covered issues of practical relevance for small and medium-sized firms like the appropriation, validation, confirmation and operationalisation of relevant knowledge.

One significant finding was that the interviewees regarded the field of Software Engineering as ill-structured and complex. This complexity could only be mastered by purposeful information-search strategies. As a main medium, the internet was said to be used due to its variety of information and speed. The search engine Google thereby obviously was a central point of departure for search activities. Though, vagueness and unreliability restrain the information value of the internet strongly.

Interestingly enough, using online forums and newsgroups instead could not overcome the disadvantages of internet use, as forums were reckoned as unstructured, complex and time-consuming, sometimes due to the perceived obligation not to use them without actively contributing to them (lurking). Forums of manufacturers (such as pages of SUN/Java), in contrast, are strongly frequented by professionals, as they do not only deliver the information needed, but additionally provide a provable decision base: the manufacturer's commitment gives the content a certain authority, as users expect him to oversee the forum and correcting wrong statements. Furthermore, the use context can be easily differentiated at a product-related base and thus problems purposefully (on the basis of the use semantics of products) be addressed.

This shows that in the German software branch, like generally in practice fields, information often is related to reputation. For example, the refusal of "too academic" knowledge is based upon the understanding that this type of knowledge mostly is presented intrinsically plausible, but, as it is impossible to implement everything appearing plausible, the relevance for the practitioners every-day work is perceived to be highly problematic. In practice, therefore, it is rather peered at "what the colleague does", or in which particular respect "the competitor does not sleep". In general, the study (Nett & Wulf 2005) showed a strong belief of the interviewees in the importance of knowledge and a strongly sceptical attitude against authoritative "academic" knowledge. Sceptical positions against authoritative knowledge systems, however, can be found in science itself.

3 KNOWLEDGE AND COMMUNITIES

Knowledge is a concept with a long philosophical tradition. In the context of the internet, knowledge often is discussed along a conception of ontologies (in plural) representing systems of formally well-defined conceptions and relations (Berners-Lee et al. 2001). This mathematical ontology model differs very much from the philosophical ontology tradition (see, e.g.: Heidegger 2005, critically: Popper 1998) with its differentiation between essence and accident, as the mathematical model does neither interpret ontology as a representation of a given, grown-up or imposed "order of things" (Foucault 1973), nor as a "Sprachspiel" (Wittgenstein 2003), but as an empirically identifiable model of knowledge representation.

Due to the lack of a uniform interpretation of the world among all individuals, generally there is no authoritative vocabulary for description and interaction (Rorty 1989). Instead in natural languages, knowledge exchange is embedded in every-day settings. When ontologies can only exist as theoretical models, knowledge can be represented differently in different real-world communities without following a thorough and formal logic. Formal models like the ontologies-conception within computer science were constructed to avoid the seemingly amorphous character of natural-language fields.

When communities were detected as a means for knowledge management, the "logical" way to deal with them seemed to construct new ones according to optimum ways of information storage and retrieval. However, communities turned out not to be merely the result of planned commitment, but of supportive environments, identities and emergent processes, as well. Therefore, communities cannot be constructed like simple machines.

The fact that Communities of Practice are not completely manageable, was often seen as a problem for knowledge management. However, the community-related, embedded nature of expertise exchange can be an advantage for human reception of knowledge, as, for example, the identity-based representation of knowledge can be used as a means to assure the reliability of sources in every-day life: our image of a speaker is influenced by how we trust him.

The Community Broker does not attempt to go beyond this natural social competence of human users, but to support it in reducing redundant procedures, which would be required in online searches otherwise. The Community Broker thus does not try to replace individual expertise on the basis of a generalized set of meta data, but to help individual expertise by automatic support of otherwise more time-consuming search activities. Its design will be presented in the following.

4 USE CASES AND USE CONTEXT

Use cases are sequences of human-computer interaction, well-known in *object-oriented software engineering* (Jacobson et al. 1992). When designing the Community Broker, use cases helped to determine the functioning of existing search tools, and to envision new related opportunities. The difference between the use context and use cases was the point of departure. The aim was to learn about the relevant use context by ethnographic field research, and to contrast this context with the use cases represented in given software. Thus new use cases could become imaginable thus allowing for conceptual reflections.

It turned out that there can be different search processes on the web. Furthermore, web searches can have different motives. For example, in the initial situation of a search process, search results can be well-defined in advance and only the information itself unknown, or one does not know exactly what to search for (cf. Saito & Omura 1998). Furthermore, the problem can be to evaluate received information respective its quality and reliability. Searching can be related to a whole set of activities of knowledge practices such as gathering, validating, recontextualising, exchanging and storing. Therefore, search activities are related to search situations and search strategies and differ considerably.

4.1 Information search on portals

Information search can follow different paths. One is to rely on the order of information pre-structured by an information supplier, like when using a portal. A user has the opportunity to enter a relevant portal and to navigate along a table of contents according to topics.

The advantage of the portal is to receive structured assistance for finding information. However, previous knowledge is a prerequisite to find an adequate portal, and to find content on it. Reading an information offer on the portal presupposes a previous appropriate search. Thus, a continuation of the search may be the result of such a search.

On the other hand, there is the necessity of information-system designers to have anticipated relevant perception and related navigation structures of the users. If this is not sufficiently given, only a tentative procedure remains for the user.

As many users look for knowledge on very concrete practice-relevant questions, but knowledge cannot be made available in such detail, abstraction is a common way to organise content. However, there can be very different ways of abstraction, and thus, of organizing information. Successful navigation does not only demand for a similar understanding, but also for a comparable application ability of the user and the information-system designer. Experience shows numerous examples indicating the problems of anticipating information demands of users, in particular, in dynamic fields of knowledge.

For some portals, the internal content can be addressed by external search engines. This is also the case of the VSEK Software Engineering portal (www.software-kompetenz.de). According to the

research on search activities of software developers, they tend to use search engines as a usual point of departure. Therefore, a plausible use case of information search on the portal can be determined.

Use Case: well-defined search of search-engines accessible portal content

The point of departure is a perceived information demand of a user motivating the use of a general search machine. Commonly known search terms are being entered into the search mask and the program starts. Among the search results, the user chooses the link to a contribution of a portal (i.e., the Software Engineering portal). By opening the link, the browser displays the related website of the portal.

A user starting with a search engine does not have the same context information like somebody, who entered the same website by starting from the starting page of the portal. This problem is reduced by the design of many portal sites, for example, by offering fix navigation bars and logos, or other hints. Nevertheless, the access of a portal via a search engine is based on similarities between the vocabulary on the site and the search terms of the search engine, whereas navigation on the portal generally follows design ideas of the information-portal designer. This can make a difference.

Problems when searching the internet are likely and common, and not even perfect portals could solve all of them satisfyingly. Therefore, the information and navigation offer of a portal cannot prevent disappointing search results. When the information displayed is not the one expected, or complicate to be understood or accessed, the user may have to react by further, often time- and resource-intensive activities. These are mostly carried out on search engines, which quite often are typical starting points of searches, in spite of the variety of information offers on portals.

Therefore, there is not only one way of search practices from search engines into portals, but, additionally, a way from portals into search engines. In search engines, search strategies usually are less influenced by anticipations of information-system designers, but based on experiences with general information search on the web. Nevertheless, there are very different problems to be tackled, as shown in the following.

4.2 Information search via general search engines

If the target search object is already well-defined, respective search terms may be entered purposefully into the search mask. If the search target, in contrast, cannot be specified at the beginning of the search (which occurs very often), the user may enter search terms into the search masks, which are vaguely considered as relevant to the goal context. To limit the search field and to evaluate links in respect to their relevance for the search object, particular links may be followed. The search could be continued entering new terms into the search mask derived from the checked pages, if they were considered useful for further search. Though, regarding the potentially high number of links, this can crop up as laborious and time-killing.

On the other hand, there is the possibility to make estimations respective to their potential relevancies for the desired target via the URLs that are connected to the links listed. However, these strategies of limitation and specification are often only insufficient auxiliary ways to be able to tackle the information overload within an appropriate time span.

Use case: web-content searched via a search engine:

The user feels an information demand and starts a search engine on his/her browser. He/She identifies search terms on the basis of assumptions. As a result, a list of links is produced, often according generic prioritizations of the search engine (e.g. frequency of calls of a website). In order to validate the use of the information for him, the user has to open the first link or choose another one (for example, due to the representation: a short text string with emphasized search terms and URL), which appears to be a plausible offer, and start to read the (sometimes very extensive) content.

In case of extensive or complicated content, the user can be forced to adopt an extensive process of excerption. This can cause the insight that the first information-search strategy has to be regarded as failed. The first search may be followed by further procedures, for instance rejecting the first search result and accessing other links, pursuing links of the first result, rejecting the entire strategy of search by entering new and different terms, or rejecting the whole search process with search engines as such.

4.3 Information search with the Community Broker

Out of the outlined use cases of established tools ([1] portals, [2] search engines), there results an option of constructing a continuative tool, being auxiliary to the former searching tools.

Use case: the Community Broker

*The user feels an information demand **when he has come to a portal website**, neither fully understanding its content nor its relevance for his/her particular situation. He/she activates the Community Broker provided on the portal. The Community Broker lists threads from diverse on-line newsgroups, where the same vocabulary is used as on the website, wherefrom the request took place. The online community threads supply diverse points of reference out of practice in the form of strings of text, contrasting to the generally described "knowledge module" of the portal. Browsing through the threads, the user decides upon his/her search strategy and either steps back to the information offer of the portal or starts a new search with his browser's or the portal's search engine.*

Reading the threads the user can decide whether the people discussing in the threads have different problems than he/she himself, or not. He/she can identify characteristics, which are similar to his/her, or differ. In any case, they have similar concepts as the abstract once on the portal website. The user thus is able to filter individual practice-relevant criteria on the basis of postings.

Therefore, the use of the Community Broker can help to decide upon the search strategy. One can go back to the portal website, study it in detail and undergo the efforts to apply the related knowledge upon the own situation. Or one may use this context knowledge to change and specify one's search strategy. As the Community Broker links the abstract models of the portal to the newsgroups which emerge out of concrete practical contexts, it provides the option of expressing individual criteria of search for unknown fields of knowledge and connecting different levels of abstraction.

By means of the contextualisation of abstract knowledge the user may develop a sense of significant criteria for his/her personal information needs. Thus, the user can be enabled to locate him-/herself within an unknown field of knowledge and to pursue the search more precisely adapted to the specific individual search target.

5 REALIZATION OF THE COMMUNITY BROKER

In the following, the realisation of the Community Broker on the basis of the use case above is been described. The Community Broker has been developed in the form of a software prototype for the VSEK Software Engineering portal (www.software-kompetenz.de). The Community Broker is a software tool which helps to find relations between online community discussion logs and websites. The bases for the relationship analysis are terms on the portal website the user visited last.

The Community Broker prototype handles only UseNet-newsgroups, which are matched against the knowledge module the portal user visited last. In principle, any type of information source which allows text-based matching calculations could be used (e.g. mailing lists, web communities, ACM.org references). The only precondition is the availability of a specific adapter, which allows access to the specific information source. Information is transformed into so-called profiles. Profiles are generic

containers, which make information independent from its sources thus enabling similarity determination. This approach of decoupling the information sources is the basis for a flexible architecture (Aier & Schönherr 2004, Krasner & Pope 1988, Stevens & Wulf 2002). The architecture of the Community Broker consists of three subsystems: Profiling, Matching and Controlling. The use of a flexible architecture facilitates the adaptability to the application domain (Krafzig et al. 2004). The sub-systems are presented briefly in the following:

- **Profiling sub-system:** The profiles represent summarized information, which is generated from heterogeneous information sources. In case of software-kompetenz.de, there are two types of profiles: knowledge-module profiles which contain information about the knowledge module selected by the user, and community profiles which are derived from newsgroup discussions. The generator component summarizes the information in four steps: filtering of "stop words" such as "the, this, and", reducing terms to word stems, statistical term analysis, and translating according to a German/English-software engineering dictionary. In addition, online-community profiles allow for meta data about the online community to be contained, like data about the number of postings per day or number of participants. Such information can be helpful for users when they want to judge the online community.
- **Matching sub-system:** When the profiles are specified, the matching subsystem calculates the degree of similarity between the created profiles, calculated from online communities and the "knowledge module" the user was on when invoking the Community Broker. The comparison algorithm used by the matching sub-system is based on the vector-space model algorithm (Baeza-Yates & Ribeiro-Neto 1999), which represents both documents and queries as vectors of terms and allows calculating similarity between them.
- **Controlling sub-system:** The controller manages the whole system and offers services to external applications (e.g. queries from users on software-kompetenz.de). In particular, the controller sub-system triggers the creation of community- and "knowledge module"-profiles. Because of performance, community profiles are created in an asynchronous manner. This means that they are created at an assigned time (e.g. over night) and not at real-time when user queries arrive.

6 CONCLUSION

Research on knowledge processes in the German software branch (Nett & Wulf 2005) showed that search activities of developers often were not based on a lack of general information, but on lacking trustworthy information. This indicates that search activities in the internet are more than the transfer of data and often related to situated learning (Suchman 1987) and expertise (Schön 1983).

When it comes to expertise for the individual, a well-structured canon of knowledge may exist with regard to core activities of an expertise field (with the "cores" differing), but additional knowledge is always necessary from heterogeneous contexts (differing also), and the field may be claimed by controversial experts. What is a canon in one field can turn out to be meaningless in another. The context of information use thus can produce a variety of valid relevant vocabularies, which can be a problem for a searcher, making it a problem to explain, what has to be searched.

However, the situated character of information demands can be an advantage instead of a problem, if the searcher is using its related knowledge in order to deploy an informed searching strategy: in heterogeneous fields, a searcher in the internet does not only search along the structures of the information required, but also along the lines of meta information on the sources, relevant discourses and actors. Such (partly subconscious) strategies are often based on identity representations in sources.

The strategic ability to evaluate vague information is based on the socialisation of the searcher, in particular, his/her situated experiences in the community structure of society, which is articulated in everyday contexts. The everyday experiences of the searcher in former Communities of Practice thus can be used later in searching activities. Searches on the internet can be supported by indicating semantic similarities between websites and forum discussions, as intended by the Community Broker.

The central topic of the paper, the importance of communities for the Community Broker, can be described by this paper only on the conceptual level, as no related practice experience and testing has been effected yet. Studies of the (hopefully) evolving use (Andriessen et al. 2003) of the Community Broker can be interesting regarding the relation between community interaction and web-search strategies.

References

- Aier, S. and Schönherr, M. (2004). Enterprise Application Integration – Serviceorientierung und nachhaltige Architekturen. Vol. 2. Gito, Berlin.
- Andriessen, J.H.E., Hettinga, M. and Wulf, V. (2003). Introduction to a Special Issue on Evolving Use of Software. *Journal of Computer-Supported Cooperative Work*, Kluwer, 12(4), 367-380.
- Baeza-Yates, R. and Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley.
- Berners-Lee, T., Hendler, J. and Lassila, O. (2001). The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities, *Scientific American*, 284(1). 34-43.
- Foucault, M. (1973). *The order of Things*. Vintage, New York.
- Heidegger, M. (2005). Gesamtausgabe, II. Abteilung: Vorlesungen 1919-1944, Band 62: Phänomenologische Interpretationen ausgewählter Abhandlungen des Aristoteles zur Ontologie und Logik. Vittorio Klostermann, Frankfurt am Main.
- Hofmann, B. and Wulf, V. (2002). Building Communities among Software Engineers: The ViSEK Approach to Intra- and Inter- Organisational Learning. In *Proceedings of the International Workshop on Learning Software Organisations* (Henninger, S. and Maurer, F. Eds.), p. 25-33, Germany, Heidelberg.
- Jacobson, I., Christerson, M., Jonsson, P. and Overgaard, G. (1992). *Object-Oriented Software Engineering: A Use Case Driven Approach*. Addison-Wesley.
- Krafzig, D., Banke, K and Slama, D. (2004). *Enterprise SOA – Service Oriented Architecture, Best Practices*. Prentice Hall PTR.
- Lave, J. and Wenger, E. (1991). *Situated Learning: Legitimate peripheral participation*. Cambridge University Press, Cambridge.
- Nett, B. and Wulf, V. (2005). Wissensprozesse in der Softwarebranche. Eine empirische Perspektive auf klein und mittelständische Unternehmen. In Gendolla, P. and Schäfer, J. (Eds.) *Wissensprozesse in der Netzwerkgesellschaft*. Transcript Verlag, Bielefeld, 147 – 168.
- Popper, K.R. (1998). *The World of Parmenides: Essays on the Pre-Socratic Enlightenment*. Routledge, London.
- Rorty, R. (1989). *Kontingenz, Ironie und Solidarität*. Suhrkamp, Frankfurt am Main.
- Saito, M. and Ohmura, K. (1998). A cognitive model for searching for ill-defined targets on the web: the relationship between search strategies and user satisfaction. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, p. 155-163.
- Schön, D.A. (1983). *The reflective Practitioner: How professionals think in action*. Basic Books, New York.
- Krasner, G.E. and Pope, S.T. (1988). A Cookbook for Using the Model-ViewController User Interface Paradigm in Smalltalk-80. *Journal of Object-Oriented Programming*, 1(3), 26-49.
- Stevens, G. and Wulf, V. (2002). Elektronische Archive in virtuellen Organisationen: Flexibilisierte Zugriffskontrolle mittels komponenten-basierter Anpassbarkeit. *Informatik-Spektrum*, 25(1), 12-21.
- Suchman, L.A. (1987). *Plans and Situated Actions*. Cambridge University Press, Cambridge.
- Wenger, E. (1998). *Communities of Practice: Learning, Meaning and Identity*. Cambridge University Press, Cambridge.
- Wittgenstein, L. (2003). *Philosophische Untersuchungen*. Suhrkamp, Frankfurt am Main.