Knowledge Management for Distributed Agile Processes: Models, Techniques, and Infrastructure

Harald Holz
University of Kaiserslautern, Germany
holz@informatik.uni-kl.de

Grigori Melnik
University of Calgary, Canada
melnik@cpsc.ucalgary.ca

Martin Schaaf
University of Hildesheim, Germany
schaaf@informatik.uni-kl.de

Abstract

This report summarizes presentations and discussions of the IEEE WETICE 2003 Workshop on Knowledge Management for Distributed Agile Processes. The main goals of the workshop were to bring together practitioners and researchers from the areas of Knowledge Management and Agile Processes from different domains to discuss the current state of ongoing research efforts and to share practical experiences with adaptation of modern Knowledge Management techniques by agile teams.

1. Introduction

Knowledge Management is currently receiving increasing attention in diverse areas such as medicine and systems engineering. Here, special focus is put on process-oriented Knowledge Management, where abstract activity descriptions serve as the primary means to capture, organize, and distribute knowledge items that are relevant during individual, actual process steps. Most approaches developed so far rely on static processes as well as on documents indexed by formalized meta-data and additional ontologies. However, these approaches are inadequate for highly dynamic and volatile processes, whose steps cannot be planned in advance, and during which new, unanticipated "knowledge needs" frequently arise. Such processes handle mostly informal documents and rely on face-to-face communication between participants. Typical examples of such processes occur in domains like medical diagnostics and disaster management.

In Software Engineering, the realization that software development processes are inherently dynamic inspired a new discipline focusing on Agile Software Development (For fundamental principles of agility, refer to Agile Manifesto [1]). These methodologies and practices embrace high rates of change. They are being increasingly applied in the industry. However, trading off explicit knowledge captured in documentation for tacit interpersonal knowledge poses new challenges, especially in the case of distributed settings, where support by proper Knowledge Management techniques is essential.

For the purpose of the workshop, the encompassing view from [2] has been adopted, where business agility is defined as "*the ability to demonstrate flexible, efficient and swift responses to changing circumstances by maximizing [the utilization of] physical and human resources.*"

Traditionally, Knowledge Management offers heavyweight techniques that require considerate effort to build and maintain the support systems. One of the reoccurring themes was whether the "marriage" of heavyweight knowledge management techniques and agile processes is needed and at all possible. In case of a positive answer, the next issue was how one remains agile while employing these techniques.

Researchers and practitioners from five countries (Canada, Germany, Italy, the Netherlands, and USA) gathered in Linz, Austria to share and discuss the ongoing research in the area of applying knowledge management techniques by agile teams.

The workshop consisted of three tracks: (1) Knowledge Management for Semantic Web; (2) Software Engineering and Knowledge Management; and (3) Process-Oriented Knowledge Management. In the following sections of the paper (Sections 2-4) we briefly describe these three sessions. Section 5 concludes this report and provides a list of open questions for future research.

2. Knowledge Management for the Semantic Web

While the amount of information available on the Internet is growing on a daily basis, and the distributed teams move towards using global knowledge repositories, it becomes increasingly more time consuming for users to find the desired information via standard keyword search, because of the large number of irrelevant hits. Semantic Web technology aims at providing applications with the

Proceedings of the Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'03)
1080-1383/03 $17.00 © 2003 IEEE

means to reason about the information contained in documents available on the Internet, e.g. retrieving only those pages that satisfy a user's current information need.

In the workshop, two contrasting approaches were discussed. In his invited talk, *H. Stuckenschmidt* [6] addressed problems of ontology-based information sharing in dynamic environments. He identified three different kinds of changes that typically occur (changing sources, changing ontologies, and new sources), and presented an approach to cope with each of these types of change.

W. Kienreich [7] presented WebRat, a tool for visualizing and refining search result sets by means of thematic landscapes. In contrast to the approach presented in [6], the system builds only on statistical properties extracted from the documents, i.e. no meta-data or lexical information is used. An evaluation of WebRat showed that the system is particularly useful for getting an overview over a domain that is new to the user, enabling her to learn the relevant keywords/concepts of this domain.

The ensuing discussions made it clear that both, heavyweight, ontology-based approaches and more lightweight, statistics-based approaches are needed in agile processes. Future research in this area should aim at developing hybrid approaches that make use of ontologies wherever available, but can also resort to statistics-based approaches. Moreover, the incremental construction and continuous evolution of ontologies, as well as the annotation of new documents supported by statistics-based approaches needs further investigation.

3. Software Engineering and Knowledge Management

The topics in this track analyzed applicability of agile software development from the knowledge management perspective.

G. Melnik [8] provided a detailed overview of knowledge sharing approaches of agile vs. traditional, plan-driven teams. In fact, the authors introduced the term "Tayloristic Methods" since the other names (traditional, rigorous, plan-driven, task-based, heavy-weight) were found to be inadequate in the course of present discussion. The authors emphasized the fact that knowledge sharing is a crucial part of both agile and Tayloristic software development processes. However, agile methods shorten the chain of knowledge transfers and potentially reduce the amount of knowledge that needs to be shared and maintained.

F. Paetsch [9] discussed the ways the requirements elicitation, analysis, and validation are performed in various agile processes. The authors compared agile development to Tayloristic (traditional) software processes as "less document-centric and more code-oriented". However, they consider this as a symptom of deeper differences: agile methods are adaptive rather than predictive and agile methods are people-oriented rather than process-oriented.

In the course of the workshop, the meaning of documentation was discussed intensively. *T. Sauer* [12] proposed a technique that allows producing certain parts of agile documentation automatically by using Event-based Design Rationales. One practical implementation of such approach was presented and the issues of managing and maintaining design rationales were examined. This paper was voted to be the best paper of the workshop.

C. A. Vissagio [10] provided recommendations on which knowledge needs are to be addressed in distributed pair programming. He reported on the results of a multi-step empirical study conducted on a group of senior undergraduate students. The results show that there is no empirical evidence that either effort or quality deteriorate significantly with distributed pairs. This confirms several findings of Stotts, Williams, Gehringer, Nagappan, Baheti, Jen and Jackson ([3][4][5]) that the code produced by distributed pairs was equal in quality to those produced both by collocated pairs and by teams not synchronously paired.

Several ideas for further experimentation in the field were discussed. For instance, a truly important study would involve agile teams maintaining the systems built by agile other teams. It would be important to assess the effort going into such maintenance work and contrast it to the effort of traditional teams. It would also be useful to experiment with the teams utilizing knowledge support tools.

4. Process-Oriented Knowledge Management

The management of knowledge is a process for itself comprising of activities like capturing, storing, and dissemination. The research for *Process-Oriented KM* (POKM) focuses on the tight integration of KM activities into the processes where knowledge is produced and consumed. Abstract activity descriptions provide necessary context information enabling new intelligent services like situation-specific knowledge dissemination. However, research for POKM has to face many new challenges. KM activities need to be minimally invasive and are frequently carried out by actors being not confident with the underlying concepts; application domains may be highly dynamic; information produced during process execution may be informal not following predefined formats. The methods discussed in this track addressed these aspects of agility, which made the application of KM support systems in many domains extremely difficult. The topics of the presentations ranged from situation-specific decision support to knowledge dissemination, weakly structured workflows, and lightweight knowledge elicitation.

H. Holz [11] presented an extension for *PRIME* (PRocess-oriented Information Resource Management Environment), a KM tool that adds proactive knowledge

delivery support to process execution systems. The authors argue that the original PRIME implementation requires some effort for setup, e.g. the formal specification of knowledge needs and potential knowledge sources. The extension utilizes collaborative filtering techniques based on Case-Based Reasoning (CBR) and, thereby, enables a heuristic-based recommendation. The authors claim that the resulting system better fits the requirements of agile teams by omitting additional knowledge engineering steps, which have to be carried out in advance, otherwise.

Proactive knowledge delivery was also a topic in the talk given by *R. Weber* [14] who discussed the *Monitored Distribution* (MD) approach for distributing lessons-learned within organizational processes. MD allows the dissemination of knowledge artifacts in a just-in-time fashion and is especially useful for achieving business agility [2].

A topic addressed by the research project *FRODO* (Framework for Distributed Organizational Memories) is the adequate workflow support for knowledge-intensive tasks. The authors argue that classical, static process models cannot describe these tasks. A solution proposed by *L. v. Elst* [15] leads to the concept of weakly structured workflows and the talk reported on the experimental evaluation of the FRODO platform for the realization of organizational memories with focus on knowledge-intensive activities. The evaluation showed the advantage of weakly structured workflows with their possibility of lazy/late modeling especially in the case of changing requirements.

M. Schaaf [13] provided an application of flexible decision support for weakly structured workflows. The authors presented the *PROGEMM* (PROcess-oriented GEneric Management of Medical Knowledge) approach that enables ad-hoc workflow configuration of clinical processes based on the current situation (diagnosis, medical treatment) and previously executed workflows. PROGEMM focuses on collaboration among physicians. According to the authors, an area which becomes more and more important due to the technological progress in health care that would require constant training of the physicians, otherwise.

F. Sartori [16] who addressed the aspect of agility for communities of practice completed the track and presented KEPT, a knowledge elicitation tool to be applied in the production process of truck tires at the Truck Business Unit of *Pirelli Tyres*. In its current implementation, KEPT enables compound designers to store their knowledge about compounds of the tire as a blend of chemical ingredients in order to obtain specific thermal-mechanical characteristics.

From the presentations and the discussions of this track can be observed that minimizing the effort for knowledge engineering activities is actually a key strategy for making POKM solutions agile. It has been achieved by using lightweight techniques like CBR, delaying the explication of knowledge to execution time, e.g. weakly structured workflows, or seamless integration of knowledge elicitation techniques. Of course, this is slightly different from the methods of newly developed agile approaches in Software Engineering where the strategy is to reduce the amount of knowledge that needs to be communicated and represented. However, purely agile methods have their limitations when it comes to the preservation of experiences and their dissemination throughout physically dispersed teams. During the execution of a process, it is extremely difficult to decide if tacit knowledge should be made explicit or not. Later on, it is often no longer possible to regain information about the situation, as it is required for representing the context of experiences. This is not acceptable especially for *knowledge intensive processes*. It became clear from the discussions with the participants working on software development, which is, of course, an inherently knowledge intensive activity, that the demand for agile POKM solutions is growing. With the approaches presented, here, steps into this direction have been made.

5. Conclusion

The talks and the discussions provided an interesting overview on research activities where the emerging need for integrating Knowledge Management and Agile Processes is tackled. In addition, we had contributions reporting on experiences from the area of *Software Engineering* where agile methods constitute a separate research discipline. The workshop attracted many people from other workshops as well and the heterogeneous backgrounds of the attendees lead to fruitful discussions. In summary, the aims of a) identifying potential synergies between *Agile Processes* and *Knowledge Management Techniques* b) discussing limitations of both approaches, and c) exploring ways for collaborative knowledge sharing in distributed teams have been achieved.

A challenge for future research will be the handling of tacit knowledge. Scalable, lightweight KM approaches requiring only minimal maintenance activities, e.g. knowledge capturing, until particular information items have been qualified, are likely to become key topics of future research. Furthermore, techniques for KM systems facilitating learning are currently emerging especially for heavyweight approaches like Ontology-based systems. Again, this can be seen as a step toward releasing users from the burden of modeling an entire domain in a consistent manner and making the systems more agile.

Acknowledgements

The workshop organizers would like to thank all workshop attendees who took part in the discussions and provided us with their thoughtful insights.

The following people participated in the program committee; we gratefully acknowledge their time and energies: Klaus-Dieter Althoff (Fraunhofer Institute for Experimental Software Engineering, Germany), Ralph Bergmann (University of Hildesheim, Institute for Mathematics und Applied Computer Science), Ansgar Bernardi (German Research Center for Artificial Intelligence GmbH, Germany), Ward Cunningham (Cunningham & Cunningham, Inc., USA), Mehmet Göker (Kaidara Software Inc., USA), Scott Henninger (University of Nebraska-Lincoln, Department of Computer Science and Engineering, USA), Frank Maurer (University of Calgary, Department of Computer Science, Canada), Charles Petrie (Stanford University, Stanford Networking Research Center, USA), Michael M. Richter (University of Kaiserslautern, Department of Computer Science, Germany), Steffen Staab (University of Karlsruhe, Institute AIFB, Germany), and Laurie Williams (North Carolina State University, Computer Science Department, USA).
In addition, we would like to thank Andrea Freßmann, Marco Spinelli (both University of Hildesheim, Institute for Mathematics und Applied Computer Science), and Prashant Baheti (North Carolina State University, Computer Science Department, USA) for contributing additional reviews.

References

[1] Agile Manifesto. http://www.agilemanifesto.org.

[2] Gartner UK Ltd.: "The Age of Agility", Report prepared by Gartner for BT, July 2002.

[3] Baheti, P., Gehringer, E., and Stotts, D., "Exploring the Efficacy of Distributed Pair Programming," Proceedings Extreme Programming/Agile Universe, Chicago, IL, 2002.

[4] Baheti, P., Williams, L., Gehringer, E., and Stotts, D., "Exploring Pair Programming in Distributed Object-Oriented Team Projects," Proceedings OOPSLA Educator's Syposium, Seattle, WA, 2002.

[5] Stotts, D., Williams, L., Nagappan, N., Baheti, P., Jen, D., Jackson, A. Virtual Teaming: Experiments and Experiences with Distributed Pair Programming. Technical Report TR03-003, Univ. of North Carolina at Chapel Hill, March 2003.

[6] Stuckenschmidt, H.: Ontologies in Dynamic Environments, in: [17].

[7] Granitzer, M., Kienreich, W., and Sabol, V.: WebRat: Supporting Agile Knowledge Retrieval through Dynamic, Incremental Clustering and Automatic Labelling of Web Search Result Sets, in: [17].

[8] Chau, Th., Maurer, F., and Melnik, G.: Knowledge Sharing: Agile Methods vs. Tayloristic Methods, in: [17].

[9] Paetsch, F., Eberlein, A., and Maurer, F.: Requirements Engineering and Agile Software Development, in: [17].

[10] Canfora, G., Cimitile, A., and Visaggio, C.A.: Lessons learned about distributed pair programming: what are the knowledge needs to address, in: [17].

[11] Holz, H., and Schäfer, J.: Collaborative, Task-Specific Information Delivery for Agile Processes, in: [17].

[12] Sauer, Th.: Using Design Rationales for Agile Documentation, in: [17].

[13] Maximini, K., and Schaaf, M.: The PROGEMM Approach For Managing Clinical Processes, in: [17].

[14] Weber, R.: Proactive Knowledge Distribution for Agile Processes, in: [17].

[15] L. van Elst, L., Aschoff, F.-R., Bernardi, A., Maus, H., and Schwarz, S.: Weakly-structured Workflows for Knowledge-intensive Tasks: An Experimental Evaluation, in: [17].

[16] Colombo, E., and Sartori, F.: KEPT: a Knowledge Elicitation Tool for supporting Communities of Practice, in: [17].

[17] Proceedings of the WETICE 2003, IEEE Press, 2003 (to appear).