Monograph of next issue (August 2008)

"EUCIP: A Model for Definition and Measurement of ICT Skills"

(The full schedule of **UP**GRADE is available at our website)

Vol. IX, issue No. 3, June 2008

Monograph: Technology-Enhanced Learning
(published jointly with Novática*)
Guest Editors: *Carlos Delgado-Kloos and Fridolin Wild*

UPENET (**UP**GRADE European NETwork)

CEPIS NEWS

* This monograph will be also published in Spanish (full version printed; summary, abstracts, and some articles online) by **Novática**, journal of the Spanish CEPIS society ATI (*Asociación de Técnicos de Informática*) at <http://www.ati.es/novatica/>.

Applying "Scruffy" Methods to Enable Work-Integrated Learning

Stefanie N. Lindstaedt, Tobias Ley, Peter Scheir, and Armin Ulbrich

This contribution introduces the concept of work-integrated learning which distinguishes itself from traditional e-Learning in that it provides learning support (i) during work task execution and tightly contextualized to the work context, (ii) within the work environment, and (iii) utilizes knowledge artefacts available within the organizational memory for learning. We argue that in order to achieve this highly flexible learning support we need to turn to "scruffy" methods (such as associative retrieval, genetic algorithms, Bayesian and other probabilistic methods) which can provide good results in the presence of uncertainty and the absence of fine-granular models. Hybrid approaches to user context determination, user profile management, and learning material identification are discussed and first results are reported.

Keywords: Associative Retrieval, Competence-Based Knowledge Space Theory, Learning On Demand, Technology-Enhanced Learning, User Context, User Profile.

1 Motivation

In current business practice and e-Learning research projects, most spending is devoted to enhancing knowledge transfer of formal training interventions. Haskell [1] informs us that in 1998 US$ 70 billion were spent on formal training and Back [2] states that in 2000 US$ 78 billion were spent on corporate training and continuing education. On the other hand, studies have revealed that in today's economy only a small amount of knowledge that is actually applied to job activities (learning transfer) comes from formal training. On average people only transfer less than 30% of what is being learned in formal training to the professional workplace in a way that enhances performance. This is independent of the kind and quality of the courses taught, but mainly depends on too little consideration of work environment needs during and after formal training efforts [3]. 80-90% of what employees know of their job, they know from informal learning [4]. Initiatives aiming at enhancing knowledge transfer of formal training try to answer the question: "How much does the learner know after engaging in

Authors

Stefanie Lindstaedt is Head of the Knowledge Services division at the Know-Center in Graz (Austria) and is scientific coordinator of the APOSDLE project. In these roles she is responsible for the management of many large, multi-firm projects and the scientific strategy of the division. Her research interests are in Knowledge Management, Technology-Enhanced Learning, and Software Engineering. For the last five years she has focused on the issue of Work-Process Integrated Learning, developing the APOSDLE concept and the concept of AD-HOC Learning. She holds a PhD and a M.S. both in Computer Science from the University of Colorado (CU) at Boulder (USA). <slind@know-center.at>.

Tobias Ley is the Deputy Division Manager and senior researcher in the Knowledge Services division at the Know-Center in Graz. Since 2001, he has managed numerous industry based research and development projects in the areas of Knowledge and Skills Management and Technology-Enhanced Learning. He is deputy to the scientific coordinator of the APOSDLE project, has developed the learning goal and competency models for this project and also leads the work package on dissemination, exploitation and standardisation. Tobias holds a PhD in Cognitive Psychology from the University of Graz. <tley@know-center.at>.

Peter Scheir is researcher and software developer at the Knowledge Services division at the Know-Center in Graz. He is responsible for the APOSDLE service-oriented software architecture and leads the work package on the integrated knowledge structure in the APOSDLE project. He has studied Telematics at the Graz University of Technology, and graduated with distinction. In addition, he is a PhD student and lecturer at the Knowledge Management Institute of the Graz University of Technology where his research interests include Concept-Based Information Retrieval and Semantic Knowledge Services. <pscheir@know-center.at>.

Armin Ulbrich works at the Knowledge Services division at the Know-Center in Graz as Head of the IT-Project Office. He has managed numerous industry based research and development projects in the areas of Knowledge Management and Context-Based Information Delivery. He holds a Diploma (M.S.) in Telematics from the Graz University of Technology. His research interests include modelling and identification of user characteristics based on user interactions, integration of user profile information from diverse sources and Social Semantic Web. <aulbrich@know-center.at>.

the formal training?" Instead, as suggested by the above numbers, the question which should be asked is: "To which extent can the learner apply the newly acquired skills to her work tasks?"

2 Work-Integrated Learning in Knowledge Work

Based on these insights, our concept of *work-integrated learning* focuses on enabling a shift from the training perspective of the organization to the learning perspective of the individual. Specifically, we are interested in exploring how informal learning happens within the work environment today and how it can be supported in the future.

The kind of on-the-job learning and learning-by-doing already takes place in companies. Otherwise people would not have been able to learn the 80-90% of the things they need to know on the job (see section 1). These questions have been addressed by research focussing on informal workplace learning [5]. However, much of what we know today is based on research in educational settings (schools and universities) or in formal workplace training. Much less research has been conducted in informal workplace learning settings.

A large scale empirical study into how people learn at the workplace has been conducted by Kooken et al. [6] who used a multi-method approach. Based on in-depth workplace observations, interviews, self reports in learning diaries and a survey conducted in knowledge intensive companies, the authors conclude that informal learning at the workplace is very frequent. It consists mainly in searching in digitally documented sources and contacting colleagues through face-to-face means. 63% and 70% of the learning episodes studied in this investigation involved these two kinds of learning strategies respectively.

With the work-integrated learning approach, we are especially seeking to support *knowledge work*. We follow Kelloway and Barling [7] when they delineate four types of knowledge work: (i) Creation of new knowledge or innovation, (ii) the application of existing knowledge to current problems, (iii) the packaging or teaching of knowledge, and (iv) the acquisition of existing knowledge through research or learning. Typical knowledge workers may include engineers, analysts, consultants, researchers and the like.

What is key, is that a person conducting knowledge work is likely to engage in *all* these activities intermittently and thereby dynamically switch to different roles in the context of their work. It is these different roles that a knowledge worker typically employs which we seek to support with the work-integrated learning approach. This is also the main distinction to more traditional (e-)Learning approaches [8]. To work, learn and teach efficiently and effectively, a knowledge worker must be provided with optimal guidance to manage the large variety of knowledge artefacts available in the corporate information infrastructure.

Based on these insights from informal learning and knowledge work research, work-integrated learning may be described by the following characteristics:

■ Individuals are responsible for their own competency development and they learn autonomously: they set their own learning goals, are responsible for time management and results, and chose their own learning strategy.

■ Individuals are enabled to learn within their own specific work processes and context: When a learning situation during work appears, the individual in general is empowered to satisfy the learning need at that time.

■ Individuals are enabled to learn within their own computational work environment: Learners are not forced to leave their familiar work environments to access learning material and contact relevant subject matter experts but are enabled to do so using their familiar tools.

■ Organizations provide the work environment (also including flexibility, time, etc.) to enable this competency advancement and actively support it in their cultures.

■ Any guidance provided must be flexible to support different knowledge work roles.

In the subsequent sections, we discuss some challenges and how these can be accommodated for by the application of advanced information and communication technologies. As a conclusion, we present concrete approaches and implementations which illustrate the use of these "scruffy" methods in organizational settings.

3 Technological Challenges

Typically e-Learning systems are a wonder of carefully designed content, fine-granular models, interdependencies and hand crafted metadata: The learning domain is broken down into meaningful learning units or modules which encompass concepts, facts and processes. They entail fine granular learning information, exercises, tests, etc. Each of these units is carefully designed using a multitude of different media appropriate for the learning type and learning purpose the unit is serving. A dependency structure identifies prerequisites and post-conditions. Based on the units, learning paths (courses) can be created by instructional designers taking into account the target group as well as preferred didactical aspects. In order to allow for improved personalization, a multitude of metadata is attached to the units. A number of metadata formats have been developed by different specification and implementation bodies, such as IMS Global Learning Consortium (IMS)[1], the Aviation Industry CBT Committee (AICC)[2] and the Dublin Core Metadata Initiative (DCMI)[3]. Designers can specify how a system should react upon certain conditions and which service facilities are to be invoked dynamically during learning (e.g. start a conference, send emails to peer learners). In addition, information on the learning resource's type of interactivity (active, expositive, mixed), semantic density (very low, low, medium, high, very high) and difficulty (very easy, easy, medium, difficult, very difficult) can be attached to learning resources.

---

[1] <http://www.imsproject.org/>
[2] <http://www.aicc.org/>
[3] <http://dublincore.org/>

In addition to metadata describing learning resources, e-Learning systems often provide detailed user models which allow for the representation of different learning levels and competency portfolios in the different areas, learning preferences, etc. A learner's peers, tutors and teachers are represented in order to allow students access to expert help.

In short, one is faced with a thoroughly designed network of interrelated pieces which need to be artfully concerted to deliver a meaningful learning experience to the user. Reflecting on these properties, one can easily understand why e-Learning content is expensive to create, requires lots of (metadata) standardization, and also requires a lot of organizational structure.

In contrast, new learning approaches such as work-integrated learning (see [8], for possible scenarios) and organizational learning put one requirement in the centre of attention: *Flexibility*. Being closer to the application of knowledge (rather than to the internalization of knowledge) such approaches critically rely on providing always the newest available content in ever changing learning situations. While in traditional course-oriented e-Learning one could still manage the large amount of design work (also because the learning domains stayed rather stable), this is not the case any more in these new settings. Here we have to strive for the best possible available learning information instead of striving for the best designed e-Learning content.

Thus, in such situations it is simply impossible to create and maintain such a carefully crafted network of interdependent learning pieces and structures. Instead, we have to move towards embracing approaches which enable us to best deal with change (while at the same time accepting their side effects such as a lower level of accuracy, likelihood of errors and not always optimal instructional design).

In the following section we discuss three functionalities crucial for the implementation of work-integrated learning: (1) user context determination, (2) user profile management, and (3) identification of relevant learning material. We present possibilities of moving away from the pure approaches of instructional design to the application of "scruffy" methods (such as associative retrieval, genetic algorithms, Bayesian and other probabilistic methods) to enable work-integrated learning. The "intelligence" within such systems may be "seen as a form of search and as such not perfectly solvable in a reasonable amount of time" [9].

4 Hybrid Approaches

The approaches discussed in the following have been developed within two ongoing projects: DYONIPS[4] (Dynamic ONtology based Integrated Process Optimisation) aims at providing a user context-aware personal information management system and APOSDLE[5] (Advanced Process-Oriented Self-Directed Learning Environment) where

the objective is to develop support tools for work-integrated learning (following the vision introduced above). From DYONIPOS we report results concerning user context determination. From APOSDLE we report results concerning user profile management and identification of relevant learning material.

The foundation of our approach (within both projects) is to not rely on specifically created (e-)Learning content, but to reuse existing (organizational) content which was not necessarily created with teaching in mind. We tap into all the resources of an organizational memory which might encompass project reports, studies, notes, intermediate results, plans, graphics, etc. as well as dedicated learning resources (if available) such as course descriptions, handouts and (e-)Learning modules. The challenge we are addressing is: How can we make this confusing mix of information accessible to the knowledge worker in a way that she can advance her competencies with it?

A frequently travelled path (also within e-Learning systems) is the creation of fine-grained semantic models which allow for the categorization and retrieval of such resources. But as we discussed above, the creation of such models, their maintenance and the annotation of resources with their concepts prove prohibitive in a dynamic environment. Thus, our approach is a hybrid one: complementing coarse grained semantic models (maintained as much as possible automatically, see below) with the power of diverse scruffy methods, improved over time through usage data and user feedback (collective intelligence).

Here the models play two roles: serving as initial retrieval triggers and providing the basis for simple inferences and heuristics to interpret user interactions. There is a trade-off for accuracy here. However, users have become increasingly accustomed to this concept through their usage of (Internet) search engines. Also, obsolete models do not provide any added value and additionally are in danger of providing a false sense of security.

The general course of action within both systems, DYONIPOS and APOSDLE is: we employ a daemon running in the background to capture and analyze a user's actions (e.g. mouse movements, keyboard entries, opening of applications). User context data thus observed, and specifically work tasks recognized, are stored in the user profiles and are improved via inference mechanisms and heuristics which interpret the user context in relationship to the task and domain models kept. This user profile data in turn is used for adapting support to the users' needs and interests. Based on the user profile data, recommendations are computed within an associative network aiming at supporting the users' learning goal attainment, the preparation of the retrieval of resources and acts of collaboration. In the following sections we describe three hybrid approaches to accomplish this.

4.1 User Context Determination

The determination of a user's work context is crucial in order to provide support for work-integrated learning.
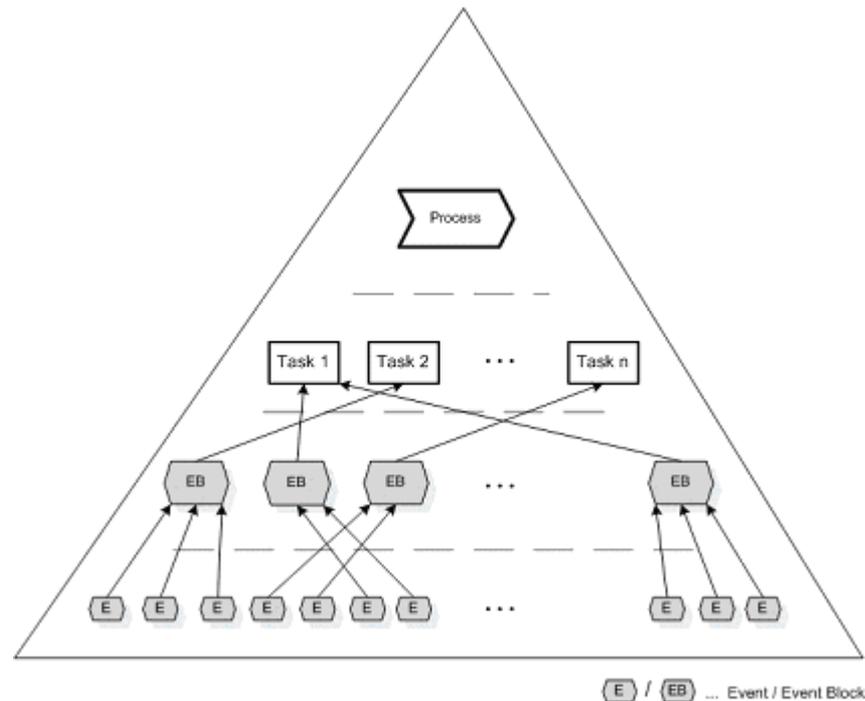

Figure 1: The Semantic Pyramid from a Knowledge Worker's Perspective Comprises the Event (E), Event Block (EB) and Task (T) Layers.

Within e-Learning systems the work context of the user is typically ignored. Instead an artificial learning context is created which is mainly based on the learning domain and learning concept to be acquired.

The DYONIPOS approach to context determination involves two phases: a training phase and a run-time phase. During the training phase task executions by a number of different users are captured and labelled. These captured execution logs are then utilized to train a classifier to distinguish between the different tasks. During run-time the context determinator continuously monitors the interactions of the user and tries to automatically classify the execution logs. If an execution log is classified with a confidence value above a certain threshold the task is recognized [10].

We use an ontology that describes the user's context by taking into account the user's interactions, the resources on which the user acts on and the corresponding (automatically extracted and manual generated) metadata relations.

The user's context model can be seen as a semantic pyramid which describes the continuous evolution of contextual information through different, semantic layers (Figure 1). Starting at the bottom with events that are executed by one knowledge worker and ending with processes where many knowledge workers can be involved.

User interactions with the system and reactions from the system to the user's interactions represent *events*. Events can be user inputs, such as mouse movements, mouse clicks, starting a program, creating a folder, a Web search, or opening a file.

An *event block* is defined by a sum of events which are chronologically ordered. An example of an event block is "editing a document on page 2". Event blocks are formed using predefined static rules, which map a set of events to an event block. Event blocks are combined into *tasks* by grouping together similar event blocks into semantic sets. The tasks obtained by this grouping are automatically learned from low-level events and event blocks which are a result of the sensor data aggregation.

The manual assignment of the event blocks to a task is used to train a classifier. The training is based on the context features we observe. Some context features which provide valuable information, e.g. the window title or the application name, are used directly for training. Other context features, e.g. user input or the content of a currently viewed document or Web page, require pre-processing (e.g., stemming, stopword removal). A detailed analysis of an extended experiment at the Austrian Ministry of Finance reports achieved accuracy of up to 75%. Here only four out of 13 context features are utilized (application name, content, window title, and semantic type) and five classification algorithms are compared.

This shows that the use of classifiers for user context determination is feasible and that the results can be the basis for user profile management and identification of relevant learning material as discussed in the next two sections.

4.2 User Profile Management
APOSDLE stores user-related context information (see

Facts about the user, e.g.
Name
Address
Email address

User's behavior during work, e.g.
Tasks executed
Collaborations participated

Further information inferred from usage data, e.g.
Potential future competency development
Potential collaboration partners

Data external to the user but still having an impact, e.g.
Domain model
Task/Competency mappings

Environment data

Inferred data
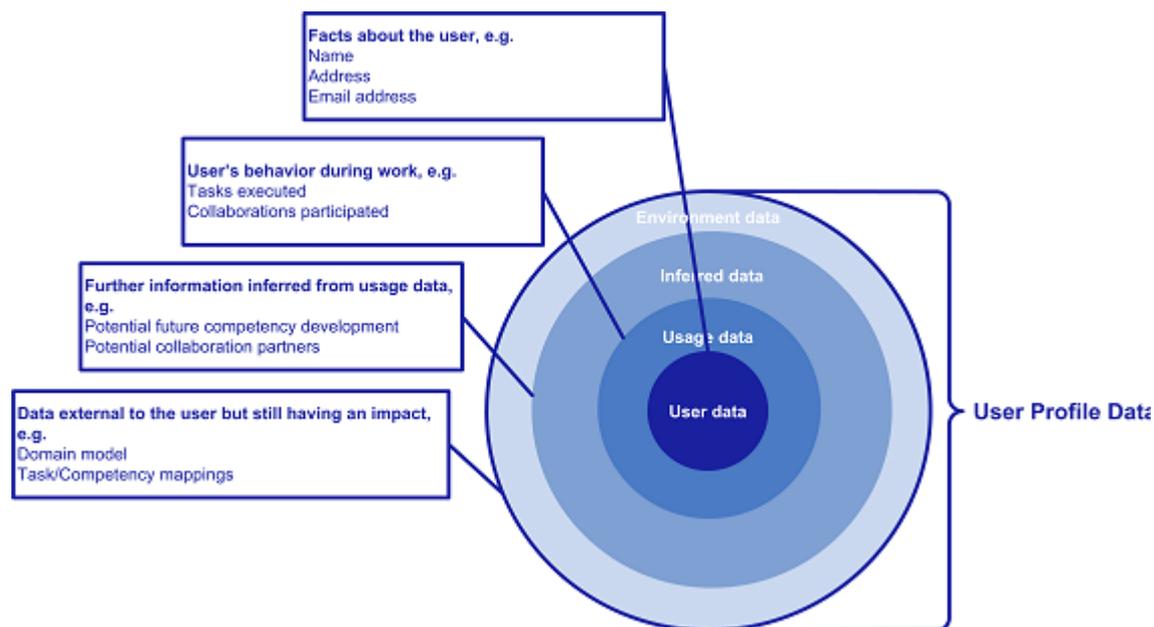
Usage data

User data

User Profile Data

Figure 2: User Profile Data Arranged in Layers as Implemented for the APOSDLE User Profile.

section above), specifically the identified tasks, in digital user profiles. These profiles are used for maintaining the user's usage history and current context with respect to their personal work, learning, and collaboration related experiences. The form chosen for representing user information has been informed by approaches from related research on user contexts (e.g. [11]) and user modelling [12]. The APOSDLE approach differentiates between four forms of user related data (see Figure 2): user data, usage data, inferred data, and environment data. This layering of user profile information allows us to clearly separate factual information and assumed information about the user. The outermost layer (environment data) is not directly related to an individual user's profile and is as such not stored within the user profile. Nevertheless, environment data has a significant impact on the user profile: The user profile services operating on the user and usage data utilize the environment data for interpretation and computing inferences.

One example of how environment data can support the interpretation of user profile data is the mapping between tasks and learning goals. A learning goal describes knowledge and skills needed to perform a task. It is defined as a discrete element of a cognitive activity (learning goal type) connected with a domain concept. The formalisms employed are based on competence-based knowledge space theory [13]. One important advantage of this theory is that it allows the computation of learning goals through a learning need analysis by comparing knowledge needed to execute a task and the knowledge state of the user. Another one is the possibility to infer a user's learning history by examining the work task she has engaged with in the past (task-based learning history).

The current APOSDLE prototype utilizes the number of work task executions as a basis for predicting which learn-

ing goals have been mastered. A preliminary simulation employing a cross validation technique resulted in moderate to high validity scores. In the future we plan to also take collaborations concerning similar topics into account.

This shows that it is feasible to use user behaviour as the basis for competency identification. Our goal is to automatically infer a variety of user characteristics through interaction analysis thus freeing the user from continually updating her user profile. Clearly user feedback plays a significant role here and we already are experimenting with different possibilities.

4.3 Identifying Learning Material

In order to provide powerful, intelligent retrieval mechanisms for work-integrated learning support the APOSDLE approach includes an associative network [14]. This associative network implements heterogeneous retrieval mechanisms: semantic retrieval (based on learning domain concepts) is seamlessly integrated with a variety of similarity-based retrieval mechanisms.

This has the advantage of on the one hand providing services with exact matched materials for instructional design tools and on the other hand also providing more inexact similarity-based services for information delivery and creativity tools. In addition, the fact that associative networks can "learn" based on changing the edge weights is used by APOSDLE to incorporate implicit as well as explicit user feedback.

The associative network relies on both, information in an ontology and the statistical information in a collection of documents. The associative network is queried by a set of concepts from the ontology and returns a set of documents. Documents in the system are (partly) annotated with ontological concepts if a document *deals with* a concept.
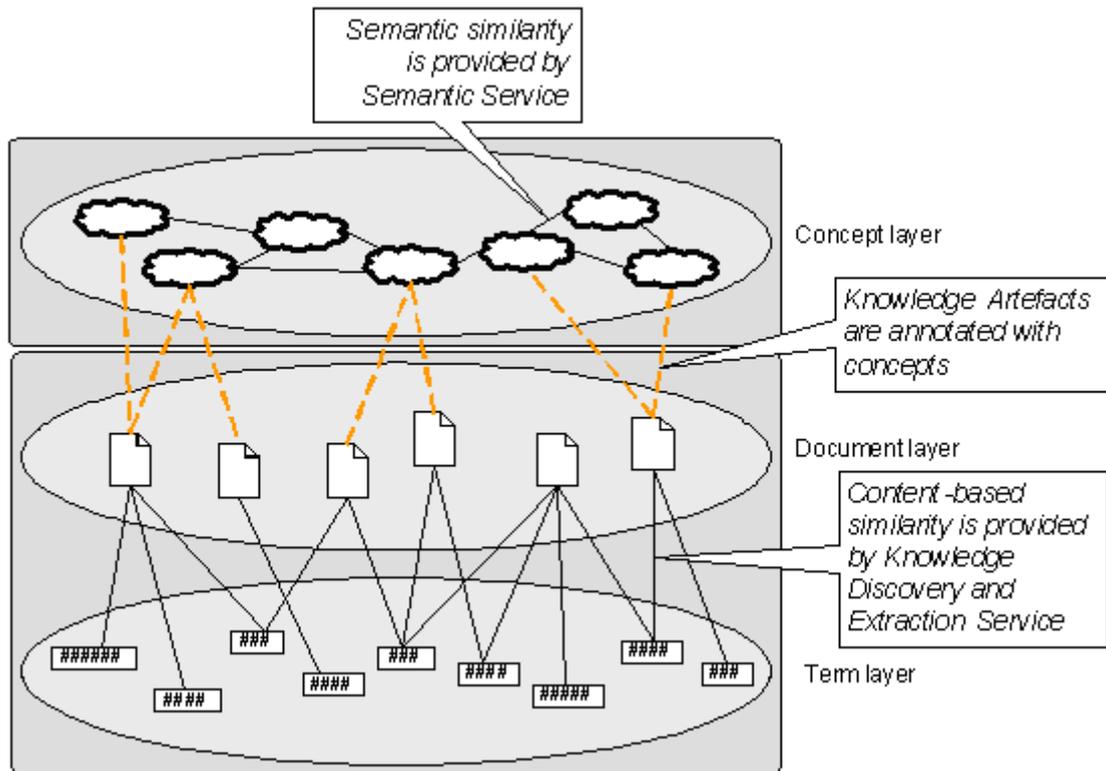
Figure 3: The Associative Network Exploits Two Different Forms of Similarity, that is Semantic Similarity and Content-Based Similarity. Semantic Annotations Function as a "glue" between these Two Forms of Similarity.

For example, if the document is an introduction to use case models it is annotated with the corresponding concept in the ontology. The annotation process is performed manually but is supported by statistical techniques (e.g. identification of frequent words in the document collection) [15].

Concepts from the ontology are used as metadata for documents in the system. Opposed to classical metadata, the ontology specifies relations between the concepts. For example, class-subclass relationships are defined as well as arbitrary semantic relations between concepts are modelled (e.g. UseCase isComposedOf Action).

The structure of the ontology can be utilized for calculating the similarity between two concepts in the ontology. This similarity can be used to extend a query by similar concepts before retrieving documents dealing with a set of concepts.

After retrieval of documents was performed, the result set can be extended by means of textual similarity. Different combinations of query and result expansion were evaluated against each other. We used data available in the first release of the APOSDLE system which was built for the domain of Requirements Engineering. The ontology contains 70 concepts and the document set consists of 1016 documents. 496 documents were annotated using one or more concepts. 21 concepts from the domain ontology were used to annotate documents. We compared eight configurations (including one baseline configuration). The results

proved encouraging since the three configurations which performed query expansion in combination with result expansion consistently performed best (see [14] for more details).

5 Conclusions

Our approach is to apply a battery of advanced "scruffy" methods to bridge the gap between coarse grained, hand-crafted semantic models and fine grained learning needs. The ultimate goal of this research is to minimize or at best fully eliminate the need for hand-crafted formal models. This will also significantly reduce the amount of human effort needed to create e-Learning systems.

Prototypes developed in the DYONIPOS and APOSDLE projects have shown that the approach is feasible, and will continue to be evaluated in detail using four different learning domains within four application partners.

References

[1] R. E. Haskell. "Reengineering Corporate Training. Intellectual Capital and Transfer of Learning". Quorum Books. 1998.

[2] A. Back, O. Bendel, D. Stoller-Schai. "*E-Learning in Unternehmen. Grundlagen - Strategien - Methoden – Technologien*". Orell Flüssli Verlag. 2001.

[3] D. G. Robinson. "Skill and Performance: They are not equal". Apartment Professional Magazine. 2003.

[4] B. Raybould. "Performance Support Engineering Part One: Key Concepts". Ariel PSE Technology, 2000. Cited in Dickover, N.T.: "The Job is the Learning Environment: Performance-Centered Learning to Support Knowledge Worker Performance", JIID, Vol.14, No. 3. 2002.

[5] M. Eraut. "Informal learning in the workplace". Studies in Continuing Education, 26 (2), 247-273. 2004.

[6] J. Kooken, T. Ley, R. de Hoog. "How Do People Learn at the Workplace? Investigating Four Workplace Learning Assumptions". In E. Duval, R. Klamma, M. Wolpers (Eds.): EC-TEL 2007, LNCS 4753, pp. 158–171. Heidelberg: Springer. 2007.

[7] E. K. Kelloway, J. Barling. "Knowledge work as organizational behavior". International Journal of Management Reviews, 2(3): 287-304. 2000.

[8] S. Lindstaedt, H. Mayer. "A Storyboard of the APOSDLE Vision", in W. Nejdl and K. Tochtermann (Eds.), "Innovative Approaches for Learning and Knowledge Sharing" (LNCS 4227), 628-633, Springer, Berlin. 2006.

[9] G. Gigerenzer, P. M. Todd. "Simple Heuristics That Make Us Smart". The ABC Research Group. Oxford University Press. 1999.

[10] A. S. Rath, N. Weber, M. Kröll, M. Granitzer, O. Dietzel, S. N. Lindstaedt. "Context-aware Knowledge Services, Personal Information Management: PIM 2008". CHI 2008 Workshop, April 5-6, 2008, Florence, Italy.

[11] A. K. Dey, G. D. Abowd, D. Salber. "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications", in: Human-Computer Interaction (HCI) Journal, 16 (2-4), 97-166. 2001.

[12] J. Fink, A. Kobsa. "User Modeling for Personalized City Tours", in: Artificial Intelligence Review, 18, 33-74. 2002.

[13] T. Ley, S. N. Lindstaedt, D. Albert. "Supporting Competency Development in Informal Workplace Learning", in K. D. Althoff, A. Dengel, R. Bergmann, M. Nick, T. Roth-Berghofer (Eds.), Professional Knowledge Management (LNCS 3782), 189-202, Springer, Berlin. 2005.

[14] P. Scheir, M. Granitzer, S. N. Lindstaedt. "Evaluation of an Information Retrieval System for the Semantic Desktop using Standard Measures from Information Retrieval". In Proceedings of Lernen-Wissen-Adaption, Halle/Saale, Germany, September 24-26, 2007, pp. 269-272. 2007.

[15] V. Pammer, P. Scheir, S. Lindstaedt. "Two Protégé plug-ins for supporting document-based ontology engineering and ontological annotation at document level". In 10th International Protégé Conference, July 15-18, 2007, Budapest, Hungary. 2007.