Chapter 10. Increasing predictability and sharing tacit knowledge in electronic design

Vadim Ermolayev[1], Frank Dengler[2], Carolina Fortuna[3], Tadej Štajner[3], Tom Bösser[4], Elke-Maria Melchior[4]

[1]Zaporozhye National University, 66 Zhukovskogo st. 69600 Zaporozhye Ukraine,
vadim@ermolayev.com
[2]Karlsruhe Institute of Technology, B. 11.40 KIT-Campus Süd D-76128 Karlsruhe Germany,
frank.dengler@kit.edu
[3]Jozef Stefan Institute, Jamova 39, SI-1000 Ljubljana Slovenia,
carolina.fortuna@ijs.si, tadej.stajner@ijs.si
[4]kea-pro, Tal, 6464 Spiringen Switzerland,
tb@keapro.net, emm@keapro.net

Abstract: The chapter reports on the use of knowledge process learning, articulation and sharing technologies developed in the ACTIVE project for increasing the performance and decreasing the ramp-up efforts of knowledge workers in engineering design projects. Attention is paid to the specific characteristics of knowledge processes in microelectronic engineering design, of which one of the most important is the absence of predefined workflows. Instead of following rigid working patterns, the knowledge workers exploit their tacit knowledge and experience for finding the most productive way through the "terrain" of the possible process continuations. The knowledge workers in this domain are design project managers, designers, and design support engineers. Process knowledge is mined from distributed heterogeneous datasets, fused, and used for visualizing the design project plan and execution information. The visualization suggests optimized performance, points to the bottlenecks in executions, and fosters collaboration in development teams. A project navigation paradigm is developed that helps knowledge workers more easily accomplish their work. We describe the software prototype architecture and implementation. Validation results are presented which indicate that the solution is helpful in providing expert assistance to design project managers performing their typical tasks of project planning and execution control.

11.1 Introduction

In knowledge-intensive sectors of industry knowledge workers (Drucker 1969) are central to an organisation's success – yet the tools they must use often stand in the way of optimising their productivity. A remedy to the defects of current knowledge worker tools has recently become substantially in demand across industries. For example in the ACTIVE project, the three case studies in consulting,

In: P. Warren et al. (eds.), Context and Semantics for Knowledge Management,
DOI 10.1007/978-3-642-19510-5_10, © Springer-Verlag Berlin Heidelberg 2011, 189-212

telecommunications and engineering design have been driven by this requirement. Knowledge workers acting alone but more importantly in teams that can be distributed geographically and organizationally are of a particular concern and focus in our research themes. One of the themes is the support for informal process knowledge acquisition, articulation and sharing.

To this theme the notion of an informal (or knowledge) process is central. The definition by Warren et al. (2009) lays out the ground for the specific features: *"Informal processes are carried out by knowledge workers with their skills, experience and knowledge, often to perform difficult tasks which require complex, informal decisions among multiple possible strategies to fulfil specific goals. In contrast to business processes which are formal, standardized, and repeatable, knowledge processes are often not even written down, let alone defined formally, vary from person to person to achieve the same objective, and are often not repeatable. Knowledge workers create informal processes on the fly in many situations of their daily work"*.

ACTIVE has adopted a service-oriented and component-based approach to its architecture. Services and components are defined at a number of levels (Warren et al. 2009). At the bottom level are infrastructure services. At the level above this, machine intelligence technology is used. For example the process mining service learns repeated sequences of action executions which constitute running processes and populates the knowledgebase with these. Finally at the top level are the applications. One of the case study applications is the management of design project (DP) knowledge in microelectronic and integrated circuit (MIC) engineering design. This case study is lead by Cadence Design Systems GmbH (www.cadence-europe.com), an engineering design services provider in this domain. It goes beyond the existing performance management solutions by providing the functionalities of the following two kinds: (i) at the back-end, the learning of design process execution knowledge from distributed datasets of acquired knowledge; and (ii) at the front-end, design project knowledge articulation and sharing – by providing a lightweight collaboration platform.

The remainder of the paper is structured as follows. Section 11.2 surveys the related work in informal process representation, mining and extraction, articulation and sharing. It also outlines the unsolved problems that are further addressed in our work. Section 11.3 presents briefly the ACTIVE approach to informal process acquisition, articulation and sharing that helps the knowledge workers in engineering design navigate their projects. Section 11.4 elaborates on the architecture and implementation of our fully functional software prototype. Section 11.5 presents the plan for and the results of the validation of the implemented software in an industrial setting. Section 11.6 discusses the results and draws some conclusions.

10.2 Related work

The research and development work presented in this chapter provided contributions in several interrelated aspects relevant to managing and productively using informal process knowledge. Our contributions implemented and integrated in the software prototype (Section 11.4) comprise: the representation of informal process knowledge in the form of ontologies; the methods for informal process mining and extraction from process logs; the methods for informal process knowledge articulation and sharing using a visualization and superimposition approach. This section analyses how our results are positioned relative to other work in these directions.

Process knowledge representation. The mainstream in process modeling is represented by enterprise and business process representations – in the form of ontologies or languages. Among the ontologies the following results have to be mentioned: the Enterprise Ontology (Uschold et al. 1998), Toronto Virtual Enterprise Ontology (TOVE) (Grüninger et al. 2000), and more recently the theoretical work by (Dietz 2006) and the reference ontology for business models developed in Interop, an EU Sixth Framework Programme (FP6) Network of Excellence (Andersson et al. 2006), see also www.interop-vlab.eu. The business process modeling community has developed a variety of languages with the major objective of representing business processes as the executable orchestrations of activities. The most prominent examples of such languages are: PSL (Bock and Grüninger 2005), BPEL and more recently WS-BPEL (docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.pdf), BPML and more recently BPDM (www.omg.org/spec/BPDM/1.0/). A more comprehensive approach to semantic business process modeling and management has been developed in the FP6 SUPER project (Hepp and Roman 2007). A major shortcoming of the listed results is that they are not supposed to provide a means to model informal processes as denoted in the definition by (Warren et al. 2009).

One of the relevant approaches to modelling and representing informal processes has been developed in the FP6 Nepomuk project (Grebner et al. 2006). A shortcoming of the process representation in Nepomuk ontologies is the limitation of the scope only to the tasks performed on the computer desktop.

Our approach to informal process representation builds on the work in dynamic engineering design process modeling of the PSI and PRODUKTIV+ projects (Ermolayev et al. 2008). Our contribution in ACTIVE lies in the development of the lightweight knowledge process representation for engineering design that is essentially a micro-ontology (Ell et al. 2010) providing a simplified yet sufficiently expressive view of a design process to be visualized for articulation and sharing. This micro-ontology is aligned with the ACTIVE Knowledge Process Model (Tilly 2010) through the PSI Upper-Level Ontology where the latter is used as a semantic bridge (Ermolayev et al. 2008).

Informal process knowledge mining and extraction. Process mining is a set of data mining techniques, focused on constructing process models out of a large number of events. The purpose of these techniques is to discover process, control, data, organizational and social structures from event logs. Practical usefulness of process mining in our setting is twofold. Firstly, it allows inferring a process model when such a model did not exist in an explicit form. Secondly, it allows devising alternative models to the primary one to enable comparison of different possible interpretations with regard to complexity and observe the extent to which the primary model is being followed. Knowing the differences between the actual process model and the mined process model is crucial when optimizing the process. In the case of microelectronic design the most interesting part of the structure is the process itself. Although process mining in general also considers more organizational and social structures (van der Aalst and Song 2004), the design process for a particular design artifact is focused mainly around a single knowledge worker and his design decisions. The knowledge process which we are trying to uncover is a product of a designer's experience and intuition and is rarely explicitly documented. This makes it valuable to ensure productivity and at the same time difficult to capture manually.

There exist several different approaches to process mining each producing models of varying expressivity. The selection of an appropriate model is influenced by the properties of the event log and the expressivity requirement of the process model. However all process models are based on the notion of states – at any point in execution the process resides in some state. Multiple-actor models permit several simultaneous states although in the electronics design domain a single process instance is usually executed by a single designer. Another differentiating point between various classes of process models is the semantics of the transition between states which affects the expressivity of the model.

In informal knowledge processes, the states are often not well-defined which requires solving this issue before tackling the process mining problem. One approach that we incorporated in knowledge process mining software from other application domains was to perform clustering on event logs and use the clusters as proxies for states (Štajner et al 2010) controlling the complexity of the process model via the desired number of clusters.

In terms of transition modeling the most straightforward approaches consider the Markovian assumption: each transition to a new state is dependent only on the previous state. Usually the transition probabilities are statistical estimates of the conditional probability of one state directly following another (Hingston 2002). We have explored this approach in related knowledge worker scenarios and discovered that simple Markovian models work well for very fine-grained low-level events (Štajner and Mladenić 2010). A side effect of using such models is that they tend to have many states and transitions which make them difficult to interpret. Because of this we often resort to de-noising the model by pruning the transitions which we consider to have little information. For this purpose Probabilistic Deterministic Finite Automata are often used, for which statistically well-founded techniques for determining significant transitions are available (Jacquemont et al. 2010).

In environments where minute variations in activity order are not critical these can be further relaxed to the conditional probability of one state following another within a time window. Such a relaxation results in a slight decrease in expressiveness since a transition only means that a particular event has occurred within a given time window before some other event. However this payoff avoids too much sparsity especially when we are constrained by having many distinct activities in a relatively short event log.

When higher expressiveness in control structures is required we can consider using the family of process mining techniques based on Petri nets implemented in the ProM framework (van Dongen et al. 2005). This approach provides for modeling patterns beyond the Markovian assumption, allowing logical structures such as conjunctions, disjunctions, splits and joins. Although the compromise is that they do not operate probabilistically an important benefit of Petri net-based approaches is that the models can also be transformed into extended Event Process Chain (eEPC) diagrams which are more familiar to process analysts and more amenable to comparison with formal process models.

All of the aforementioned models can be expressed with particular subsets of PSI ontology terms. In that sense the PSI Suite of Ontologies provides the common knowledge representation formalism for knowledge integration, fusion and visualization.

Informal process knowledge articulation and sharing. . The spiral of knowledge (Nonaka and Takeuchi, 1995) introduces different knowledge conversions which is a fundamental part of sharing knowledge. People can share tacit knowledge with each other (socialization), but this is a rather limited form of sharing knowledge. Knowledge articulation within companies is the process of making tacit knowledge explicit (externalization). This explicit knowledge can be combined with other explicit knowledge (combination) and shared throughout an organization. Other employees extend and reframe their tacit knowledge with explicit knowledge by internalizing it (internalization). There are different ways to articulate and share informal process knowledge, but in all cases the informal process knowledge has to be made explicit. For instance process knowledge can be visualized manually or with tool support. In contrast to the visualization approach, the process knowledge can also be stored and shared within the system by using it directly for recommendations (Dorn et al. 2010).

Articulation and sharing using visualization approach. It is natural for a human to use visualized representations of artifacts in general and of processes in particular. Research in psychology, human memory models, image recognition and perception reveals that graphical representations are comprehended much easier and with lower effort than equivalent textual ones (Crapo et al. 2000). Therefore process visualization is one of the mature instruments to articulate processes thus enabling users to easily understand the logic of a process.

Most process visualization techniques are included in process modeling activities, which can be centralized or decentralized. An abundance of modeling methods and tools like ARIS (Scheer and Jost 2002) and IDEF3 (Mayer et al. 1995) have been developed to ease the standardization, storage, and sharing of process visualization. Unfortunately these tools are not sufficient for modeling collaborative, decentralized processes. Therefore other approaches like CPM (Ryu and Yücesan 2007) have been introduced.

In the area of knowledge processes additional methods and tools like KMDL (Gronau 2005), PROMOTE (Woitsch and Karagiannis 2005) and CommonKADS (Schreiber et al. 1999) have been developed extending the methods and tools mentioned above. In addition, semantic wikis combine the collaborative aspects of wikis (Leuf and Cunningham 2001) with Semantic Web technology to enable large-scale and inter-departmental collaboration on knowledge structures. Such features of semantic wikis have been extended to support process development (Dengler et al. 2009), enterprise modelling (Ghidini et al. 2009) and workflows (Dello et al. 2008).

Our contribution in process visualization is the enhancement of the existing Semantic MediaWiki (SMW) process development approach (Dengler et al. 2009) to visualize and discuss informal processes.

Our project navigation approach is based on offering a collaboration platform to knowledge workers that facilitates socializing, externalizing and internalizing design project knowledge using visualization. Visualizaton of project knowledge helps to combine and internalize explicit project knowledge in a way that suggests productive continuations of project execution.

10.3 The approach to project knowledge navigation

The goal of the presented case study is providing a software tool for design project managers in MIC that will articulate and facilitate sharing knowledge about good development practices in this domain.

An objective of a project manager as a knowledge worker is *finding a reasonable balance between the available and the achievable in order to meet the requirements of a customer and accomplish development in his project with the highest possible productivity*. The complexity of this task in modern design environments is beyond the analytical capabilities of even an experienced individual. A manager has to find an optimum in a solution space that has many facets: product structure comprising possibilities for block reuse; the compositions of the development team involving required roles and capabilities of the available individuals; the choices of the tools for performing design and corresponding design methodologies; the resources available for the project; project constraints and business policies; etc. One more complication may appear in the course of the execution of the project – the circumstances may change because of external events. Hence a previously good plan may turn out to be not acceptable for the follow-up. Re-planning may therefore be required at any moment.

Project managers use their working experience and intuition for taking planning decisions under these complex conditions. In fact they rely on following good practices and using the suggested development methodologies that they used in the past and which constitute their *tacit* working knowledge of project management. Our working hypothesis in this research was that offering a software tool for:

- Eliciting good development practices as stable working patterns from the design project logs
- Visualizing those practices of past projects, the plan and the state of the execution of a current project
- Facilitating moderated discussions among the members of the development team on different aspects of a project

– will decrease the complexity of making decisions for a project manager and increase the robustness of his knowledge work. Such a tool would essentially make the *tacit* knowledge of project managers within a company *explicit* – i.e. articulate and facilitate sharing good project management and engineering design practices.

For checking this hypothesis the software tool prototype of a Design Project Visualizer has been developed in the case study. The tool implements a project navigation metaphor – helping a knowledge worker find a productive execution path through the state space of an engineering design project.

It is known for informal processes in general and the processes of engineering design in particular that the paths to a desired outcome can not be specified in advance – before the process starts. Instead, a knowledge worker has to make his decision about a follow-up action by choosing among the possible continuation alternatives in an arbitrary process state – very similarly to the decisions made by a driver on the road. Drivers use navigation systems that suggest the ways to go for bypassing traffic jams, choosing a faster or a cheaper way. A similar approach is employed in our work for helping a project manager to make decisions about the continuations of his design project.

The Design Project Visualizer, like a car navigation system, provides the visualized views of the basic “terrain” map. These views are product structures, methodology flows that are either generic or bound to a particular product structure, Work Breakdown Structures (WBS). These representations are essentially provided by a project manager in a top-down fashion when he plans and kicks-off the project.

The Design Project Visualizer also assists in finding out where the project is on the “terrain” at a specific point in time. The knowledge about the execution of the project is mined from the available project log datasets, transformed to the terms of the used ontology, stored to the knowledgebase, and superimposed onto the project execution plans.

Unlike a car navigation system the Design Project Visualizer is a tool for team work. It provides the infrastructure and the functionality for moderated discussions attached to a visualized representation of any kind of a project constituent. By that it facilitates making more informed decisions that are also more transparent to the team members and are elaborated and approved with their active participation.

For constructing the necessary building blocks of the project maps and execution tracks we first looked at the tasks of the project managers in their everyday work and extracted the typical tasks of the project planning and execution management that may be effectively facilitated. Those typical user tasks (Fig. 11.1) are:

- Analyze the requirements and develop the structure of the product
- Choose development methodology and compose the team
- Develop the work breakdown structure
- Monitor the execution of the project

After extracting the typical tasks we decided about the requirements for the functionality of the software tool. The requirements were elaborated by looking at the working practices of a project manager in MIC design and extracting the use case scenarios (Ermolayev et al. 2010).

10.4 Prototype architecture and implementation

The architecture of the fully functional prototype of the ACTIVE Design Project Visualizer is pictured in Fig. 10.1. It comprises both the back-end and the front-end components and involves several ACTIVE technologies. Design process knowledge acquisition is done mainly at the back-end while the functions of knowledge articulation and sharing are offered by the front-end. As shown in Fig. 10.1, the prototype helps users to perform their typical tasks of design project management. Therefore it could be classified as a project management tool. The tool monitors the environments of the managed engineering design project, that are design systems, and allows for run-time extraction of the process-related knowledge in a bottom-up fashion. The normative, methodological and static parts of project knowledge are provided via the external tools in a top-down manner. The external tools are the Cadence ProjectNavigator and the Cadence Flow Infrastructure (CFI) Framework – c.f. Fig. 10.1. Hence, the prototype exploits the superimposition of the top-down and bottom-up project knowledge for making its articulation and sharing more efficient and effective.

Acquisition is done by incremental collection of the new knowledge about the executions of design processes through monitoring design processes in their environments and mining the dataset containing design process execution logs – using the ACTIVE Process Mining component based on the probabilistic temporal process model (TNT) (Grobelnik et al. 2009). The approach to process mining is based on the generation of the Hidden Markov Models (Rabiner 1990). As outlined in Fig. 10.1 the ACTIVated Design System Framework tools monitor the design environments and the design processes and collect the data about the low level events in the respective datasets. The datasets are further fed to the Process Mining Service of the ACTIVE Knowledge WorkSpace (AKWS) Server that produces the instances of the segments of the executed design processes in terms of the PSI Suite of Ontologies. These instances are further stored in the Cadence Knowledgebase.

Articulation and sharing are done by visualizing different facets of DP knowledge in the collaborative front-end using the SMW (Krötzsch et al. 2007) as a platform – the ACTIVE DP Visualizer. Visualization functionality is structured around the typical tasks that DP managers perform in their everyday business (upper part of Fig. 10.1). The kinds of visualization pages are those for: product structures; generic methodologies; product-bound methodologies[1]; tools; and actor roles. These primary functionalities are supported by decision making instruments for conducting moderated discussions – the discussion component as an extension of SMW and LiveNetLife (www.livenetlife.com), an application for contextualized interactive real-time communication between the users of a web site.

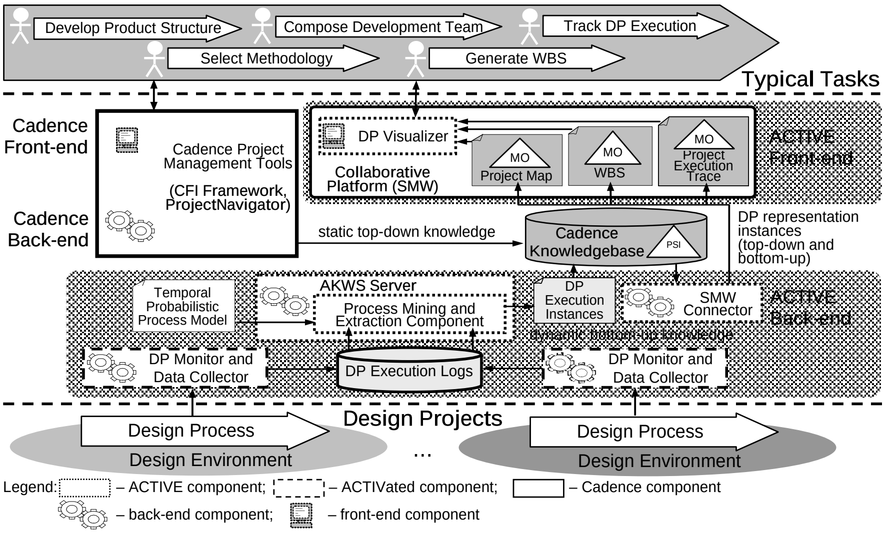


Fig. 10.1 The configuration of ACTIVE and Cadence technology components for design project knowledge acquisition, articulation and sharing.

10.4.1 Knowledge representation

A challenge in the development of knowledge representation for the case study and the software prototype of the Design Project Navigator was finding a proper balance between:

- The background result, the PSI Suite of Ontologies for MIC Engineering Design domain used at Cadence, and the model of a knowledge process (KPM) developed in ACTIVE (Tilly 2010)
- The expressive power of the knowledge representation of the Cadence knowledge base (PSI Ontologies) and the lightweight character of the enterprise

---

[1] A product bound methodology is a superimposition of the segments of the generic methodologies appropriate for the particular types of functional blocks in the structure of the product to be designed.

knowledge structures developed in ACTIVE caused by the lightweight character of the SMW used for the prototype development

The first aspect was essentially a harmonization problem. For harmonizing the KPM with the PSI Suite of Ontologies the PSI Upper Level ontology has been used as a semantic bridge (Ermolayev et al. 2010b). Please refer to isrg.kit.znu.edu.ua/ontodocwiki/ for the online documentation of this suite of ontologies. The harmonization process was bidirectional. On one hand the Suite of PSI Ontologies has been refined by cleaning the representations of process patterns and processes. This work led to fixing the v.2.3 release of the PSI Suite. On the other hand the KPM has been revised by aligning it to the PSI-ULO. This work led to the final release of the KPM (Tilly 2010).

The second problem was the selection of the minimal required part of the PSI Core Ontologies v.2.3 as the lightweight background knowledge representation for the Design Project Visualizer. For that the requirements based on the analysis of the typical user tasks and use cases have been applied and resulted in the development of the micro ontology for the case study (Ell et al. 2010).

10.4.2 The back-end: process mining and extraction

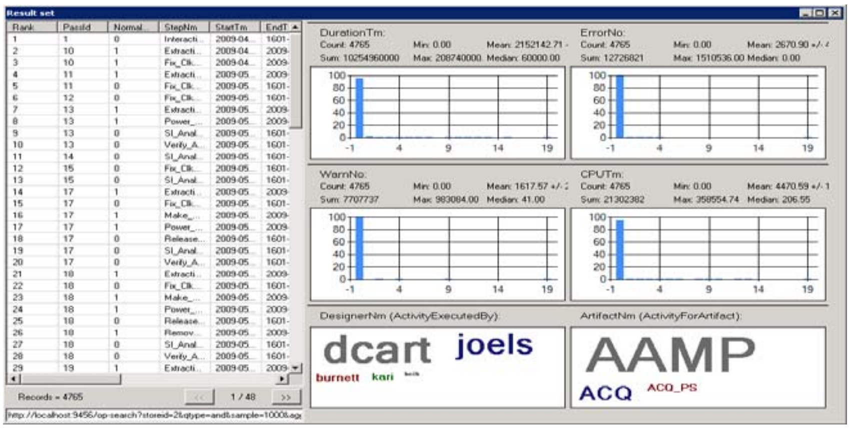
The actual implemented workflow is as follows: first, the designer's workstation is instrumented with logging tools that capture his activities, their outcome and measure the time that was required to complete those activities. Once the data is exported the process mining software loads the sequence logs and constructs a process model based on probabilistic deterministic finite automata.

The fitness for inclusion of a transition between two generic tasks in the process is evaluated using the following procedure: given an error rate α_1 and sample size, we use a statistical sequence mining technique to determine constraints for inclusion of individual transitions in the process model as presented in (Jacquemont et al. 2009). We apply a criterion called proportion constraint. Given a desired risk factor and an actual process execution log, we can compute the empirical probability of every possible transition from one state to another. This can then be used as a basis for determining whether every transition in the process is statistically significant given the observed grounding in the process execution log. A benefit of using a statistical approach is that the only parameter that the process analyst needs to specify is the risk factor which corresponds to the expected false positive rate. This parameter is easier to understand and specify than some arbitrary probability threshold. We have found that pruning the model following this approach does not affect the predictive power too adversely, while significantly reduces complexity. Following that, the software outputs the two sets of results:

- The process pattern model expressing process state transition patterns as Generic Tasks and Generic Activities – action patterns. The model also specifies statistical dependencies between individual action patterns as possible output and input configurations.

- The actual design process instances in terms of Tasks, States and Activities to which concrete Actor and Design Artifact representation instances are related. An instance of the whole design process is considered a top-level Task managed by the Actor. The top-level Task comprises the lower-level Tasks for each pass of the design process. The passes are further decomposed into atomic steps that are represented as leaf-level Tasks and Activities. After the execution of each of these steps the resulting Design Artifact Representation becomes different reflecting the fact that each step of the process brings the process closer to the target final Representation – the tape-out of the Chip.

This output is the native input for the visualization and sharing infrastructure, provided by the back-end functionality (SMW Connector, Fig. 10.1). At the same time, this provides a common interface for the consumption of process models and instances from other PSI-based design project and process management tools. From the perspective of process instances the solution resembles to some extent the ProM framework which defines the MXML format for expressing event logs. However ProM does not prescribe any terms for expressing process models. We observe that the use of the ontology, PSI Suite in our case, for expressing both process models and process execution logs is preferable in terms of integration for the purpose of informal process knowledge management.



Rank	PassId	Normal...	StepNm	StartTm	EndT
1	1	0	Interacti...	2009-04...	1601-
2	10	1	Extracti...	2009-04...	2009-
3	10	1	Fix_Clk...	2009-04...	2009-
4	11	1	Extracti...	2009-05...	2009-
5	11	0	Fix_Clk...	2009-05...	1601-
6	12	0	Fix_Clk...	2009-05...	1601-
7	13	1	Extracti...	2009-05...	2009-
8	13	1	Power_...	2009-05...	2009-
9	13	0	SI_Anal...	2009-05...	1601-
10	13	0	Verify_A...	2009-05...	1601-
11	14	0	SI_Anal...	2009-05...	1601-
12	15	0	Fix_Clk...	2009-05...	1601-
13	15	0	SI_Anal...	2009-05...	1601-
14	17	1	Extracti...	2009-05...	2009-
15	17	0	Fix_Clk...	2009-05...	1601-
16	17	1	Make_...	2009-05...	2009-
17	17	1	Power_...	2009-05...	2009-
18	17	0	Release...	2009-05...	1601-
19	17	0	SI_Anal...	2009-05...	1601-
20	17	0	Verify_A...	2009-05...	1601-
21	18	1	Extracti...	2009-05...	2009-
22	18	0	Fix_Clk...	2009-05...	1601-
23	18	1	Make_...	2009-05...	2009-
24	18	1	Power_...	2009-05...	2009-
25	18	0	Release...	2009-05...	1601-
26	18	1	Remov...	2009-05...	2009-
27	18	0	SI_Anal...	2009-05...	1601-
28	18	0	Verify_A...	2009-05...	1601-
29	19	1	Extracti...	2009-05...	2009-

Records = 4765 1 / 48

http://localhost:9456/op-search?storeid=2&qtype=and&sample=1000&ag

DurationTm:
Count: 4765 Min: 0.00 Mean: 2152142.71 -
Sum: 10254960000 Max: 208740000 Median: 60000.00

ErrorNo:
Count: 4765 Min: 0.00 Mean: 2670.90 +/- 4
Sum: 12726821 Max: 1510536.00 Median: 0.00

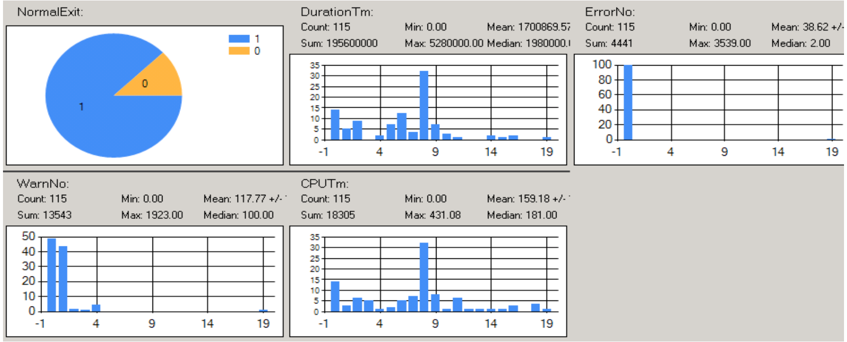
WarnNo:
Count: 4765 Min: 0.00 Mean: 1617.57 +/- 2
Sum: 7707737 Max: 983084.00 Median: 41.00

CPUTm:
Count: 4765 Min: 0.00 Mean: 4470.59 +/- 1
Sum: 21302382 Max: 358554.74 Median: 206.55

Fig. 10.2 Designers, design artifacts and development activities mined from the US dataset.

Furthermore, the Miner tool allows more complex queries of the mined design process steps, among which also queries related to times spent on particular steps, iterations, etc. For instance, it can be seen which designers executed a high number of activities (dcart in the example in Fig. 10.2) and wich artifact required these activities (AAMP in Fig. 10.2). The tabular interface allows browsing for more details in the results of a query – please refer to Fig. 10.2. Other information such as CPU time, duration of time spent by the designers per artifact or per step is also possible to query and visualize using the Miner interface. By checking the numbers for errors logged in the design process we observe that the median value was zero and the average value was around 2600. However the maximum number of logged errors was over 1.5 million per activity while the sum was almost 13 million errors. The tool allows tracking down the project that generated these errors and the designer that was executing it. At a closer inspection, it can be seen that these errors occured during one iteration of the design of NV_RAIL. Details about the most frequent activities as well as their sequence executed within that design process can be visualized. These visualizations may further be used for taking remedial actions in a design process.

In the example dataset the most frequent activity turns out to be Extraction_PRHO. This activity has normal accomplishments (there was no forced manual or automatic abortion), on average it takes about 30 minutes to complete (1700869 miliseconds), low CPU time and low number of errors/warnings (Fig. 10.3). However, the second most frequent activity – SI_Analysis_Signoff – generally does not have a normal termination, on average demonstrates more errors and warnings, and on average takes over 3 hours to be completed (Fig. 10.4). Due to the abnormal exists the duration of this activity is often not known (i.e. end time is not logged). The abnormal exits are correlated with unknown durations of activities.



NormalExit:

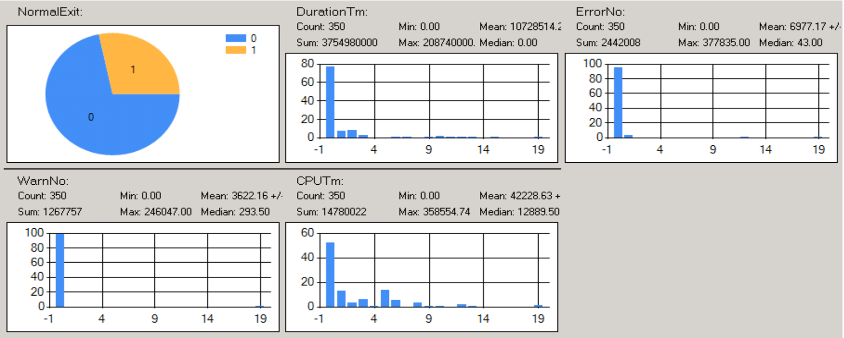
DurationTm:
Count: 115 Min: 0.00 Mean: 1700869.57
Sum: 195600000 Max: 5280000.00 Median: 1980000.

ErrorNo:
Count: 115 Min: 0.00 Mean: 38.62 +/-
Sum: 4441 Max: 3539.00 Median: 2.00

WarnNo:
Count: 115 Min: 0.00 Mean: 117.77 +/-
Sum: 13543 Max: 1923.00 Median: 100.00

CPUTm:
Count: 115 Min: 0.00 Mean: 159.18 +/-
Sum: 18305 Max: 431.08 Median: 181.00

Fig. 10.3 Details for the Extraction_PRHO activity.



NormalExit:

DurationTm:
Count: 350 Min: 0.00 Mean: 10728514.2
Sum: 3754980000 Max: 208740000. Median: 0.00

ErrorNo:
Count: 350 Min: 0.00 Mean: 6977.17 +/-
Sum: 2442008 Max: 377835.00 Median: 43.00

WarnNo:
Count: 350 Min: 0.00 Mean: 3622.16 +/-
Sum: 1267757 Max: 246047.00 Median: 293.50

CPUTm:
Count: 350 Min: 0.00 Mean: 42228.63 +
Sum: 14780022 Max: 358554.74 Median: 12889.50

Fig. 10.4 Details for the SI_Analysis_Signoff activity.

10.4.3 The front-end: process visualization and discussion

The front-end – ACTIVE DP Visualizer – for process visualisation and discussion of design projects is based on SMW and has been implemented by extending the existing process visualization approach (Dengler et al. 2009) and by developing additional result printers for SMW to visualize and export the WBS – namely the Gantt chart and the XML export result printers that are the part of the Mediawiki Semantic Project Management extension: (www.mediawiki.org/wiki/Extension:Semantic_Project_Management). The screenshots of the characteristic features of the DP Visualizer are shown in Fig. 10.5.

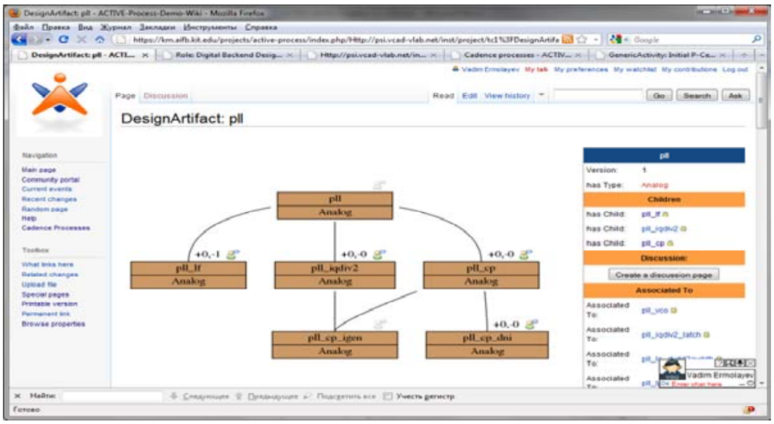
This extension builds on the capability to query for semantic properties which is provided by SMW and displays query results as process graphs, Gantt charts or XML-files containing the WBS in XML schema to be further imported into MS-Project. For the back-end a software connector has been developed that imports the knowledge stored in the Cadence Knowledgebase into the SMW pages (Fig. 10.1).

Each element of the micro-ontology (Ell et al., 2010) is represented as a wiki page containing annotated links to other wiki pages (properties). These properties are queried via an inline query language provided by SMW and the result is rendered into the destination format required by the different visualization libraries.

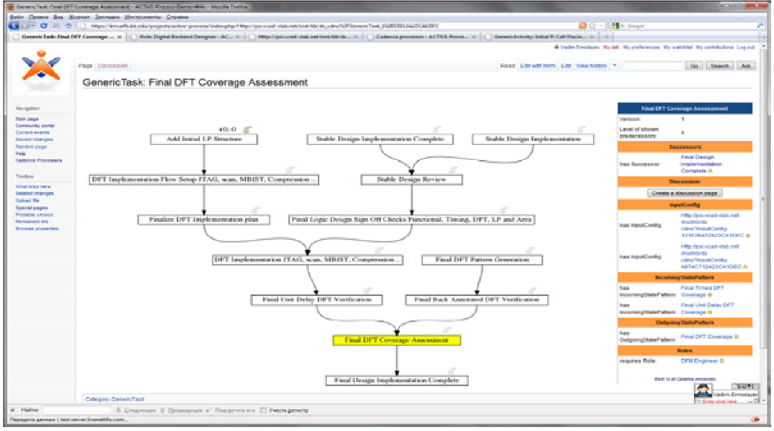
For supporting collaboration the functionality for working with talk pages has been developed as another SMW extension (Fig. 10.6). Talk pages corresponding to visualized project elements can be created to discuss pros and cons of product structures, methodologies, WBS and project execution progress. This collaborative discussion functionality has been enhanced with semantics to add metadata to each comment and allow querying. Therefore special wiki templates (www.mediawiki.org/wiki/Help:Templates) and semantic forms (www.mediawiki.org/wiki/Extension:Semantic_Forms) have been developed. A summary icon is used to display corresponding discussion activities including the sum of pros and cons (Fig. 10.5) within the product structure, methodology, and Gantt chart visualizations.

10.5 Validation Setup and Results

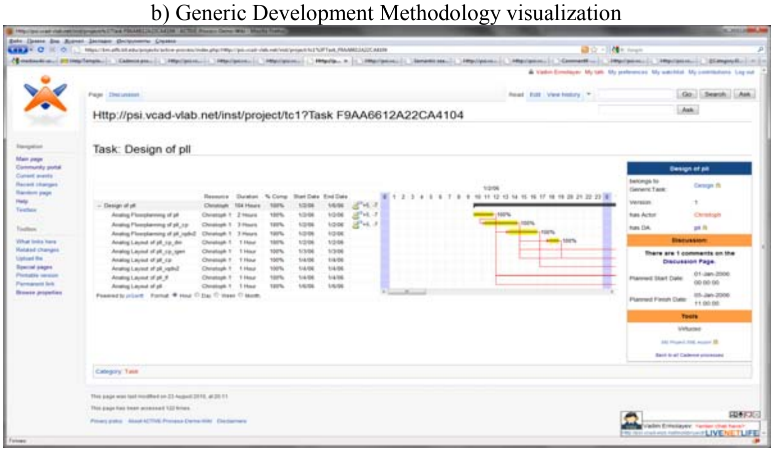
The development of the fully functional prototype was conducted in an iterative design process focussed strongly on user needs and on the organizational requirements for the application. The chip design process is very demanding in terms of adherence to detailed technical requirements and standards, consequently we have to focus strongly on testing detailed user and design processes. Different types of tests were carried out throughout the development process (Melchior and Bösser 2011). Each software version was tested repeatedly in a sequence of tests,



a) Product Structure visualization

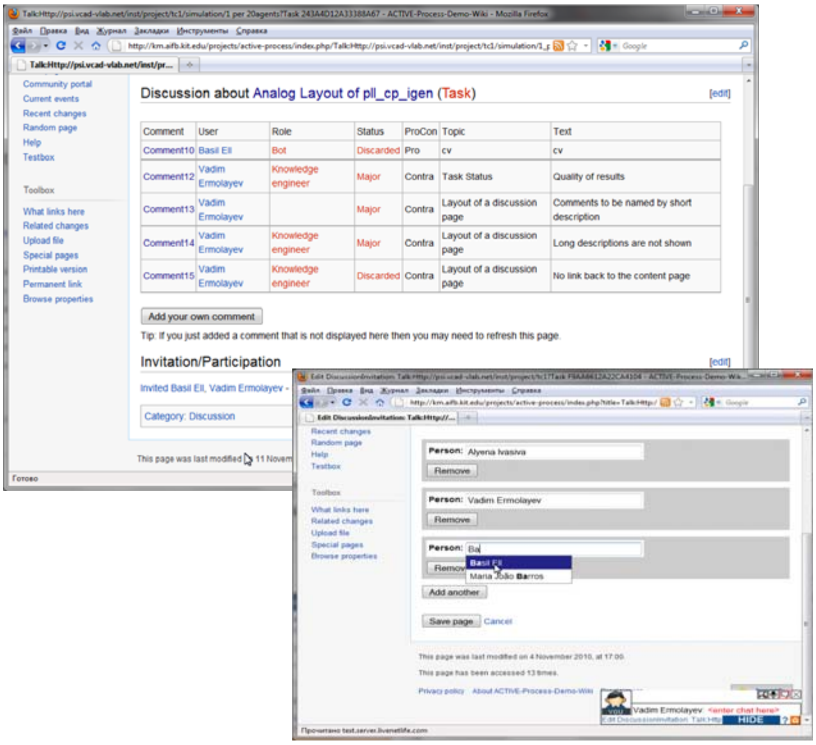


b) Generic Development Methodology visualization



c) Work Breakdown Structure with superimposed execution status

Fig. 10.5 Characteristic features of the Design Project Visualizer.



Discussion about Analog Layout of pll_cp_igen (Task)

Comment	User	Role	Status	ProCon	Topic	Text
Comment10	Basil Ell	Bot	Discarded	Pro	cv	cv
Comment12	Vadim Ermolayev	Knowledge engineer	Major	Contra	Task Status	Quality of results
Comment13	Vadim Ermolayev		Major	Contra	Layout of a discussion page	Comments to be named by short description
Comment14	Vadim Ermolayev	Knowledge engineer	Major	Contra	Layout of a discussion page	Long descriptions are not shown
Comment15	Vadim Ermolayev	Knowledge engineer	Discarded	Contra	Layout of a discussion page	No link back to the content page

Tip: If you just added a comment that is not displayed here then you may need to refresh this page.

Invitation/Participation

Invited Basil Ell, Vadim Ermolayev -

Category: Discussion

This page was last modified 11 Novem

Person: Alyena Ivasiva

Person: Vadim Ermolayev

Person: Ba

This page was last modified on 4 November 2010, at 17:00.

This page has been accessed 13 times.

Fig. 10.6 Collaborative discussion functionality for managing invitations and working with arguments

where as a general rule the subsequent level of test was carried out after the lower level test returned a satisfactory result, usually after a number of iterations:

- T1. Dry Runs and Technical Appropriateness: Experts and user representatives who are familiar with the application context assess if the software is bug-free and consistent with the requirements (Ermolayev et al. 2010), and suggest design improvements.
- T2. Usability in representative user tasks: User representatives and experts assess the software in representative tasks (Section 10.3) according to quality-of-use criteria and conformance to requirements.
- T3a. Information quality for users: Experienced users assess the quality of information generated by the application.
- T3b. User satisfaction, acceptance: Representative samples of users use the application in a realistic working context. User satisfaction and indicators for the likely acceptance of the application are measured.

The validation results described here are part of T2 of the final application prototype, with a main objective to assure that the software is satisfactory and acceptable as a tool for the typical tasks of the prospective user population. The likely effect on productivity will be considered, but this is the main question to be answered in the subsequent T3 test that is reported in (Ermolayev et al. 2011).

10.5.1 Validation Plan

The validation trials were planned in two phases and are defined by combinations of: (i) a software component; (ii) a representative user task used as a frame for the validation; (iii) an external tool used as a benchmark for the assessment; (iv) a source dataset. The summary of the validation plan is given in Table 10.1.

Table 10.1 Validation phases and types for different validated components

Validated components	Dry Runs (**T1**)	Usability for typical tasks (**T2**)	Information Quality for Users (**T3a**)		User Satisfaction, Acceptance of solutions (**T3b**)	Comment
			Completeness	Correctness		
Phase 1: Validation based on the Simulated Data. Verification tool - ProjectNavigator						
Back-end						
DPE	-	-	-	-	-	
Front-end						A trialist assesses the components by performing a typical task on a simulated project.
SS	+	+	*	*	-	
PV	+	*	*	*	-	
DC	+	+	*	*	-	
LNL	+	+	+	+	-	
Phase 2: Validation based on the US Dataset. Verification tool – CFI Framework						
Back-end						A trialist assesses the components by performing a typical task on a real project that has been accomplished and logged in the past.
DPE	+	-	+	+	+	
Front-end						
SS	-	+	+	+	+	
PV	-	+	+	+	+	
DC	-	-	+	+	+	
LNL	-	-	+	+	+	

Legend: DPE – Design Process Miner and Instance Extractor; SS – Semantic Search component; PV – Design Project Visualizer; DC – Discussion component; LNL – LiveNetLife component; “-“ – not validated; “*” – partially validated because the data is simplified and artificial (simulated project); “+” – validated.

Generic validation workflows have been developed for all the kinds of validation trials (example in Fig. 10.7). These workflows, though identical in their nature and goals, differ in the use of validation metrics, instruments (kinds of the questionnaires to be filled in) and collaboration patterns – some are for individual execution while the other are for a group of collaborating trialists.

Within the validation phases the set of validation tasks have been specified based on the typical user tasks. The generic workflows have been instantiated for each validation task. For example the task of validating usability (T2, Fig. 10.7) has been decomposed into four lower level tasks as pictured in Fig. 10.8. In turn each of these lower-level tasks have been performed using different instantiations of the generic validation workflow for T2 developed according to the validation scenarios of the particular lower level tasks. One example is given in Fig. 10.9.

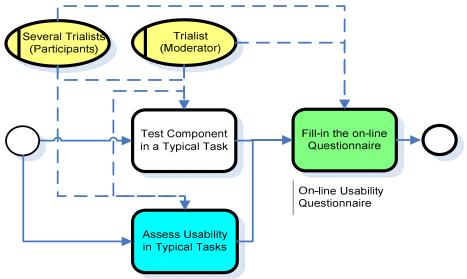


Fig. 10.7 The collaborative generic workflow for usability (T2) validation trial.

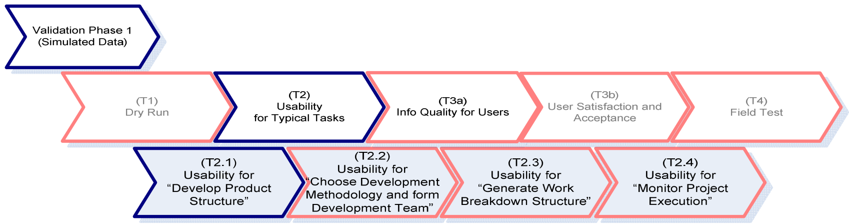


Fig. 10.8 The hierarchy of validation tasks for usability validation (T2) at Phase 1. The lowest level corresponds to the typical user task based validation.

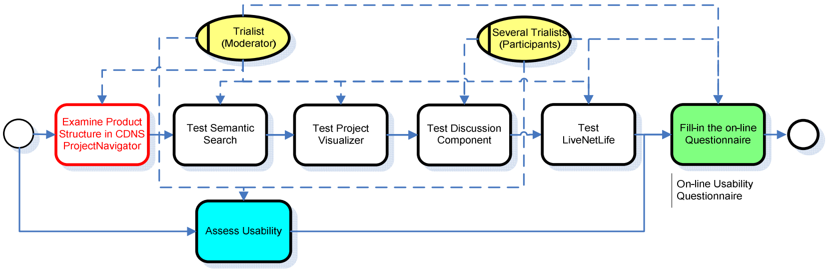


Fig. 10.9 The instantiation of the generic validation workflow (T2) for the validation task T2.1 highlighted in Fig. 10.8.

The summary of the front-end component validation plan at phase 1 is given in Table 10.2.

Table 10.2 Validated front-end components per a typical user task.

Typical user task	Semantic Search (**SS**)	Project Visualizer (**PV**)	Discussion Component (**DC**)	Live-NetLife (**LNL**)
Front-end				
Develop Product Structure	*	*	+	+
Choose Development Methodology and form Development Team	*	*	+	+
Develop Work Breakdown Structure for the project	-	*	+	-
Monitor the execution of the project	-	*	+	-

The expert test persons either decided if specific requirements are met or not met (yes, undecided, no, unable to answer question), or made quality assessment on a 7-valued scale. Three separate on-line questionnaires with 75 questions in total were answered by each test person.

10.5.2 Validation procedure

The objectives of the validation was to verify that the Design Project Visualizer corresponds to the specified technical and functional requirements of users, and to validate the quality of use of the prototype with focus on the appropriateness of the functionality for the tasks of professional users.

The test was performed by six professional experts with different roles in Cadence: Engineering director with profound expertise in design, verification and implementation; project manager; design project manager; knowledge engineer.

The evaluation was based on a task scenario composed of four typical tasks with several sub-tasks each (described above). The test users performed the tasks on their own; the discussion task was carried out in cooperation with the other test users. After the completion of each task an on-line checklist for testing conformance related to the task was completed. After executing the entire task scenario each expert completed an on-line questionnaire with questions about the quality of use of the prototype and the usefulness of the functionality for the task scenario. The three questionnaires were composed of standard and proven scales, and questions related to the specific functionality and context of the chip design process.

The critical question to be answered is whether the prototype is sufficiently mature for further, full scale tests. The experts testing the prototype are representative of the decision makers who will decide the acceptance, and thus their decision determines the organizational acceptance of the solution.

10.5.3 Validation results

Conformance with technical requirements. In total, 6 senior experts participated in the test. The resulting sample size was therefore too small for a statistical analysis. Hence, we discussed the results on a case-by-case basis putting focus on the justifications and explanations given by these experts in their assessments of the prototype. Overall, five of six professional experts have accepted and approved the Quality of the Design Project Visualizer, and four experts approved the Completeness of the Design Project Visualizer. The following reservations were formulated:

One expert did not approve the Quality of the prototype because the software was not able ”to fully visualize the complex structures, interfaces and correlations of Design Projects.” Two experts did not approve the Completeness of the prototype because of a “lack of flexibility, agility and completeness” of design project visualizations.

Several experts were undecided about the quality and correctness of the Design Project Visualizer because some specific elements (for example interfaces between the functional blocks within a design artifact) while considered essential were not included in the simulated input data. Therefore quality and correctness could not be proven in this respect. For instance four experts were undecided whether the Gantt chart representation of Work Breakdown Structures indicates the progress of the project appropriately.

Several inconsistencies were detected in the representation of design project elements and described in technical detail in (Melchior and Bösser 2011).

Conformance with functional requirements for the Discussion Component. The majority of experts (4 out of 5) agree that the functionality of the Discussion Component meets their requirements. One expert was unable to subscribe to discussions and thus did not exercise all of the functionality. Some experts are undecided about the functionality for summary boxes because the summary boxes contained summaries for simulated design project data only. This may be an issue for further investigation, or at least a further test with real data.

Quality of Use of the Design Project Visualizer. The opinions about the visual presentations of product, generic methodology and Work Breakdown Structure descriptions are divided, ranging from somewhat positive to somewhat negative. All experts but one disagree or are undecided that the visual presentations are appropriate for performing the typical tasks of the task scenario.

The information quality of the visualizations was doubted: “… a visual representation can only represent a part or a high level view of the overall process or working patterns”. The visualization of working patterns, dependencies, roles, tools etc are “too fragmented and difficult to connect for a non-savvy project manager”. “... the WBS or Gantt representation can not capture the full content and properties which are needed to perform a design activity” .

To conclude, the experts have raised doubts about the quality of use of the Design Project Visualizer.

Quality of Use of the LiveNetLife component. The quality of use of the LiveNetLife component is judged positively. However, LiveNetLife is competing with tools which are currently in use at Cadence. LiveNetLife would have to demonstrate a differentiating benefit.

It was observed that LiveNetLife does not compute the similarity with other users very reliably, a feature which might be an added value over competing tools.

Conclusions. The results of the validation of the Design Project Visualizer indicate that the solution can provide expert assistance to design project managers performing the typical tasks of project planning and execution management. According to the professional experts the conformance with requirements tested on the basis of simulated data in the Design Project Visualizer meet the technical and functional requirements. The critical statements of experts relate to added value. This must be proven by the information quality, which is the innovative feature enabled by semantic backend processes.

The reservations may also be due to inconsistencies in the Knowledge Base. The next test will be conducted with a real data set. Apart from checking the information quality the objective of this test will be evaluating if the prototype can handle large and complex projects (scalability) – another important issue which must be investigated further using real data.

Overall, after repeated iterations the components of the prototype have achieved a mature level. The functionality for discussions meets the requirements and is judged to be satisfactory for users.

The quality of use of the prototype overall fulfills minimum usability requirements, "though some features could have been implemented in a more functional and user friendly way. The reason is the lightweight nature of the basic platform (SMW)". There is potential to improve semantic search, although users can cope with the shortcomings of semantic search because the navigation and browsing in the SMW works well.

The added value of an innovative application is important for its acceptance and uptake. Diverse expert opinions about the added value of the Design Project Visualizer can be explained by the different roles of the experts. Engineers prefer to keep administrative work at a minimum level and therefore do not recognize the added value of the tool directly. Project managers currently collect administrative information manually (for example in project meetings). Automated data collection and representation in Gantt charts would be an added value for this group of users. The experts see the basic user functionality as acceptable but remain to be persuaded of the benefit of the new technology. Some experts asked how the Design Project Visualizer will improve the productivity.

The prototype was compared with competing tools used at Cadence (e.g. the visualization of Functional Blocks, re-used IPs, IPs libraries and interfaces are already captured in existing Cadence tools). The experts fear that the integration of a Design Project Visualizer into their work processes may cause additional overhead (e.g. by having to ensure the consistency of several databases) instead of increasing productivity. High upfront cost for individual users incurs a substantial lag before benefits and added value are visible. Therefore users have to be convinced that improved information quality will offset the upfront cost. The added semantic backend-functionality should add significant value for users by providing them with information with higher (pragmatic) information quality. This should improve the cost / benefit ratio sufficiently to assure acceptance by the organization and individual users at the workplace. Further tests will be conducted to collect data on this issue.

10.6 Summary

The chapter presented the results of the case study of the ACTIVE integrated project on the use of knowledge process learning, articulation and sharing technologies for increasing the performance and decreasing the ramp-up efforts of knowledge workers managing designs of Microelectronics and Integrated Circuits. One of the most important characteristics of a design project in general and in this domain in particular is a very low proportion of the use of predefined workflows. Due to that the processes in engineering design are to a substantial extent informal. Instead of following rigid working patterns, the knowledge workers exploit their tacit knowledge and experience for finding the most productive way through the “terrain” of possible process continuations. Design product structure and methodology knowledge is collected from the project manager and the members of the development team in a top-down manner. Design process execution knowledge is mined from process log datasets in a bottom-up fashion, fused, superimposed on the top-down knowledge, and further used for visualizing the design project plan and execution information in a way that suggests optimized performance, points to the bottlenecks in executions, and fosters collaboration in development teams. A project navigation paradigm has been developed that helps knowledge workers more easily find their way to a reliable outcome. This approach has been implemented in a software prototype – the Design Project Visualizer. The first results of the validation of the software prototype indicate that the solution is helpful in providing expert assistance to design project managers performing their typical tasks of project planning and execution management. The total cost / benefit improvement remains to be vindicated taking into account both organizational objectives and the fact that for some users, notably design engineers, additional overhead may be created by the tools.

Acknowledgement

The research leading to these results has been funded in part by the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement IST-2007-215040.

References

Andersson B, Bergholtz M, Edirisuriya A, Ilayperuma T, Johannesson P, Gordijn J, Grégoire B, Schmitt M, Dubois E, Abels S, Hahn A, Wangler B, Weigand H (2006) Towards a reference ontology for business models. In: ER'06: Proceedings of the 25th International Conference on Conceptual Modeling, Springer-Verlag, Berlin Heidelberg

Bock C, Grüninger M (2005) PSL: a semantic domain for flow models. Software Systems Modeling 4(2): 209-231

Crapo A W, Waisel L B, Wallace W A, Willemain T R (2000) Visualization and the process of modeling: a cognitive-theoretic view. In: KDD '00: Proc. 6th ACM SIGKDD international conference on knowledge discovery and data mining, ACM, New York

Dello K, Nixon L, Tolksdorf R (2008) Extending the Makna Semantic Wiki to support workflows. In: Proceedings of the 3rd Semantic Wiki Workshop

Dietz J L G (2006) Enterprise ontology. Theory and methodology. Springer Verlag, Berlin

Dengler F, Lamparter S, Hefke M, Abecker A (2009) Collaborative process development using Semantic MediaWiki. In: Proceedings of the 5th Conference of Professional Knowledge Management, Solothurn, Switzerland

Dorn C, Burkhart T, Werth D, Dustdar S (2010) Self-adjusting recommendations for people-driven ad-hoc processes. In: Hull R, Mendling J, Tai S (eds.) Business Process Management. Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg

Drucker P F (1969) The age of discontinuity: guidelines to our changing society. Heinemann, London

Ell B, Dengler F, Djordjevic D, Garzotto F, Krötzsch M, Siorpaes K, Vrandecic D, Wögler S (2010) Conceptual models for enterprise knowledge – final models and evaluation. ACTIVE Project

Ermolayev V, Keberle N, Matzke W-E (2008) An upper level ontological model for engineering design performance domain. In: ER'08: Proceedings of the 27th International Conference on Conceptual Modeling, Springer-Verlag, Berlin Heidelberg

Ermolayev V, Jentzsch E, Fortuna C (2010a) Requirements for fully functional prototype. ACTIVE Project

Ermolayev V, Ruiz C, Tilly M, Jentzsch E, Gomez-Perez J M (2010b) A context model for knowledge workers. In: Proceedings of the 2nd International Workshop on Context, Information and Ontologies (CIAO 2010), http://ceur-ws.org/

Ermolayev V, Groos P, Fortuna C, Štajner T, Dengler F., Ell B, Melchior E-M (2011) Fully functional prototype and validation report. ACTIVE Project

Ghidini C, Kump B, Lindstaedt S, Mahbub N, Pammer V, Rospocher M, and Serafini L (2009) Moki: the enterprise modelling wiki. In: The Semantic Web: Research and Applications. Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg

Grebner O, Ong E, Riss U, Brunzel M, Bernardi A, Roth-Berghofer T (2006) Task management model. NEPOMUK Project

Grobelnik M, Mladenić D, Ferlež J (2009) Probabilistic temporal process model for knowledge processes: handling a stream of linked text. Conf. on Data Mining and Data Warehouses (SiKDD 2009), Ljubljana, Slovenia

Gronau N, Müller C, Korf R (2005) KMDL – capturing, analysing and improving knowledge-intensive business processes. J Universal Computer Science 11(1): 452-472. doi: 10.3217/jucs-011-04-0452

Grüninger M, Atefy K, Fox M S (2000) Ontologies to support process integration in enterprise engineering. Computational & Mathematical Organization Theory 6(4): 381-394. doi: 10.1023/A:1009610430261

Hepp M, Roman D (2007) An ontology framework for semantic business process management. In: Oberweis A, Weinhardt C., Gimpel H, Koschmider A, Pankratius V, Schmizler B (Eds.): eOrganisation: Service-, Prozess, Market-Engineering, 1, Universitaetsverlag Karlsruhe

Hingston P (2002) Using finite state automata for sequence mining. Australian Computer Science Communications 24(1): 105-110. doi: 10.1145/563857.563814

Jacquemont S, Jacquenet F, Sebban M (2009) Mining probabilistic automata: a statistical view of sequential pattern mining. Machine Learning 75(1): 91-127. doi: 10.1007/s10994-008-5098-y

Krötzsch M, Vrandecic D, Völkel M, Haller H, Studer R (2007) Semantic Wikipedia. J of Web Semantics 5(4): 251-261. doi: 10.1016/j.websem.2007.09.001

Leuf B, Cunningham W (2001) The Wiki way: collaboration and sharing on the Internet. Addison-Wesley

Mayer R, Menzel C, Painter M, Witte P D, Blinn T, Perakath B (1995) Information integration for concurrent engineering (IICE) IDEF3 process description capture method report, Knowledge Based Systems Inc.

Melchior E-M, Bösser T (2011) Results of the field trials: User acceptance, business benefits, and organizational impact of ACTIVE technology. ACTIVE Project

Nonaka I, Takeuchi H (1995) The knowledge-creating company. Oxford University Press

Rabiner L R (1990) A tutorial on hidden Markov models and selected applications in speech recognition. In: Readings in speech recognition, Morgan Kaufmann Publishers Inc., San Francisco

Ryu K, Yücesan E (2007) CPM: A collaborative process modeling for cooperative manufacturers. Advanced Engineering Informatics 21(2): 231-239. doi: 10.1016/j.aei.2006.05.003

Scheer A-W, Jost W (2002) ARIS in der Praxis, Springer Verlag, Berlin

Schreiber G, Akkermans H, Anjewierden A, de Hoog R, Shadbolt N, van de Velde W, Wielinga B (1999) Knowledge engineering and management: the CommonKADS methodology. MIT Press

Štajner T, Mladenić D (2010) Modeling knowledge worker activity. In: Proceedings of the Workshop on Applications of Pattern Analysis, Cumberland Lodge

Štajner T, Mladenić D, Grobelnik M (2010) Exploring contexts and actions in knowledge processes. In: Proceedings of the 2nd International Workshop on Context, Information and Ontologies (CIAO 2010), http://ceur-ws.org/

Tilly M (2010) Dynamic models for knowledge processes. Final Models. ACTIVE Project

Uschold M, King M, Moralee S, Zorgios Y (1998) The enterprise ontology. The Knowledge Engineering Review 13(1): 31-89. doi: 10.1017/S0269888998001088

van der Aalst W M P, Song M (2004) Mining social networks: uncovering interaction patterns in business processes. In: Proc Int Conf on Business Process Management (BPM 2004), Lecture Notes in Computer Science, Springer-Verlag, Berlin

van Dongen B F, de Medeiros A K A, Verbeek H M W, Weijters A, van der Aalst W M P (2005) The ProM framework: a new era in process mining tool support. In: Application and Theory of Petri Nets, Springer-Verlag, Berlin Heidelberg

Warren P, Kings N, Thurlow I, Davies J, Bürger T, Simperl E, Ruiz Moreno C, Gòmez-Pèrez J M, Ermolayev V, Ghani R, Tilly M, Bösser T, Imtiaz A (2009) Improving knowledge worker productivity – the ACTIVE integrated approach. BT Technology Journal 26(2): 165-176

Woitsch R, Karagiannis D (2005) Process oriented knowledge management: a service based approach. J Universal Computer Science 11(4): 565-588. doi: 10.3217/jucs-011-04-056