Björn Johansson
Bo Andersson
Nicklas Holmberg (Eds.)

LNBIP 194

Perspectives in Business Informatics Research

13th International Conference, BIR 2014
Lund, Sweden, September 22–24, 2014
Proceedings

BUSINESS INFORMATICS BIR RESEARCH

Springer

Lecture Notes in Business Information Processing 194

Björn Johansson
Bo Andersson
Nicklas Holmberg (Eds.)

Perspectives in Business Informatics Research

13th International Conference, BIR 2014
Lund, Sweden, September 22-24, 2014
Proceedings

Springer

Volume Editors

Björn Johansson
Bo Andersson
Nicklas Holmberg
Lund School of Economics and Management
Department of Informatics
Lund, Sweden
E-mail: bjorn.johansson@ics.lu.se
E-mail: bo.andersson@ics.lu.se
E-mail: nicklas.holmberg@ics.lu.se

ISSN 1865-1348 e-ISSN 1865-1356
ISBN 978-3-319-11369-2 e-ISBN 978-3-319-11370-8
DOI 10.1007/978-3-319-11370-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014948359

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Business Informatics is a discipline that combines Information and Communication Technology (ICT) with the knowledge of management. It is concerned with the development, use, application, and the role of management information systems and other ICT solutions. It is an established academic and research discipline. This is evidenced by the fact that many universities offer degrees in business informatics. The academic teaching programs are not detached from the research. The area of research that lies in the focus of business informatics is constantly evolving. It is clearly visible when you look at the topics of subsequent conferences under the eternal name Perspectives in Business Informatics Research (BIR). The BIR conference series was established 14 years ago as the result of a collaboration of researchers from Swedish and German universities. The goal was to create a forum where researchers in business informatics, both senior and junior, could meet and discuss with each other. The conference series has a Steering Committee, to which one or two persons from every appointed organizer are invited. So far, BIR conferences were held in: Rostock (Germany – in 2000, 2004, 2010), Berlin (Germany – 2003), Skövde (Sweden – 2005), Kaunas (Lithuania – 2006), Tampere (Finland – 2007), Gdańsk (Poland – 2008), Kristianstad (Sweden – 2009), Riga (Latvia – 2011), Nizhny Novgorod (Russia – 2012) and Warsaw (Poland 2013). This year's 13th International Conference on Perspectives in Business Informatics Research (BIR) was held in Lund (Sweden), during September 22rd and 24th, 2014, at the Lund School of Economics and Management, part of Lund University.

This year the BIR conference attracted 71 submissions from 14 countries: ranging from Norway to Brazil, and from Russia to Portugal. They were rigorously reviewed by 55 members of the Program Committee representing 22 countries. As the result, 27 full papers, presenting novel research results were included in the proceeding.

The papers presented at the conference cover many aspects of the business informatics research, and this year there is a particular emphasis on business process management, enterprise and knowledge architecture, information systems and services, organizations and information systems development. Apart from the main conference satellite events workshops and a Doctoral Consortium were held during the first day of the conference.

We would like to thank everyone who contributed to the BIR 2014 conference. First of all we thank the authors being ready to present their research, we appreciate invaluable contribution from the members of the Program Committee and external reviewers and we thank all the members of the local organization team from the Department of Informatics at Lund University for their help in the organization of the conference. We acknowledge the EasyChair development team for providing a convenient tool for preparing the proceeding and the Springer

publishing team for their collaboration. We would also like to acknowledge IBM for their interest in being a sponsor for BIR 2014 and all the input provided from IBM making the conference such a successful event. Last but not the least, we thank the Steering Committee and we hope that BIR 2014 will be a memorable link in the BIR conference series.

July 2014 Björn Johansson

Organization

Program Chair

Björn Johansson	Lund University, Sweden

Program Co-chairs

Bo Andersson	Lund University, Sweden
Nicklas Holmberg	Lund University, Sweden

Program Committee

Jan Aidemark	Linneaus University, Sweden
Esma Aimeur	University of Montreal, Canada
Eduard Babkin	State University (Nizhny Novgorod), Russia
Per Backlund	University of Skövde, Sweden
Janis Barzdins	University of Latvia, Latvia
Rimantas Butleris	Kaunas University of Technology, Lithuania
Sven Carlsson	Lund University, Sweden
Fergal Carton	University College Cork, Ireland
Rogerio Carvalho	Universidade Federal Fluminense, Brazil
Witold Chmielarz	Warsaw University, Poland
Jan Devos	Ghent University, Belgium
Niclas Eberhagen	Linnaeus University, Sweden
Peter Forbrig	University of Rostock, Germany
Bogdan Ghilic-Micu	Bucharest University of Economic Studies, Romania
Janis Grabis	Riga Technical University, Latvia
Michal Gregus	Comenius University, Slovakia
Horst Günther	University of Rostock, Germany
Markus Helfert	Dublin City University, Ireland
Amin Jalali	Stockholm University, Sweden
Kalinka Kaloyanova	University of Sofia, Bulgaria
Christina Keller	Jönköping International Business School, Sweden
Marite Kirikova	Riga Technical University, Latvia
Andrzej Kobylinski	Warsaw School of Economics, Poland
Vladimir Krylov	MeraLabs, Russia
Markus Lahtinen	Lund University, Sweden
Michael Le Duc	Mälardalen University, Sweden
Barbara Livieri	University of Salento, Italy

Ulf Melin	Linköping University, Sweden
Mirella Muhic	Lund University, Sweden
Charles Møller	Aalborg University, Denmark
Lina Nemuraite	Kaunas University of Technology, Lithuania
Jyrki Nummenmaa	University of Tampere, Finland
Jacob Nørbjerg	Aalborg University, Denmark
Enn Õunapuu	Tallinn University of Technology, Estonia
Wipawee Paulsson	Lund University, Sweden
Tomas Pitner	Masaryk University, Czech Republic
Nava Pliskin	Ben-Gurion University of the Negev, Israel
Vaclav Repa	University of Economics, Czech Republic
Pedro Ruivo	Universidade Nova de Lisboa, Portugal
Lazar Rusu	Stockholm University, Sweden
Kurt Sandkuhl	The University of Rostock, Germany
Andrzej Sobczak	Warsaw School of Economics, Poland
Piotr Soja	Cracow University of Economics, Poland
Chris Stary	Johannes Kepler University, Austria
Janis Stirna	Stockholm University, Sweden
Eva Söderström	University of Skövde, Sweden
Bernhard Thalheim	Christian Albrechts University Kiel, Germany
Olgerta Tona	Lund University, Sweden
Benkt Wangler	University of Skövde, Sweden
Anna Wingkvist	Linnaeus University, Sweden
Stanislaw Wrycza	University of Gdansk, Poland
Styliani Zafeiropoulou	Lund University, Sweden
Jelena Zdravkovic	Stockholm University, Sweden
Iryna Zolotaryova	Kharkiv National University of Economics, Ukraine

Additional Reviewers

Anna-Sofia Alklind Taylor	University of Skövde, Sweden
Osama Mansour	Linneaus University, Sweden
Grażyna Paliwoda-Pękosz	Cracow University of Economics, Poland
Przemysław Polak	Warsaw School of Economics, Poland
Ofira Shmueli	Ben-Gurion University of the Negev, Israel
Per Svejvig	Aarhus University, Denmark
Torben Tambo	Aarhus University, Denmark

BIR Steering Committee

Kurt Sandkuhl	University of Rostock, Germany (Chair)
Eduard Babkin	State University (Nizhny Novgorod), Russia
Per Backlund	University of Skövde, Sweden

Rimantas Butleris	Kaunas University of Technology, Lithuania
Sven Carlsson	Lund University, Sweden
Peter Forbrig	University of Rostock, Germany
Horst Günther	University of Rostock, Germany
Björn Johansson	Lund University, Sweden
Mārīte Kirikova	Riga Technical University, Latvia
Harald Kjellin	Kristianstad University College, Sweden
Andrzej Kobyliński	Warsaw School of Economics, Poland
Lina Nemuraite	Kaunas University of Technology, Lithuania
Jyrki Nummenmaa	University of Tampere, Finland
Eva Söderström	University of Skövde, Sweden
Bernd Viehweger	Humboldt University, Germany
Benkt Wangler	University of Skövde, Sweden
Stanisław Wrycza	University of Gdańsk, Poland

LUND UNIVERSITY

School of Economics and Management

Table of Contents

Business, People, and Systems

Business and Information Systems Development

Contextualized Evaluation of Business Informatics

Capability Modeling: Initial Experiences

George Bravos[1], Tania González[2], Jānis Grabis[3], Martin Henkel[4], Lauma Jokste[3], Hasan Koc[5], and Janis Stirna[4]

[1] FreshTL, United Kingdom
gebravos@hua.gr
[2] EVERIS, Spain
tania.gonzales.cardona@everis.com
[3] Information Technology Institute, Riga Technical University, Kalku 1, Riga, Latvia
{grabis,lauma.jokste}@rtu.lv
[4] Department of Computer and System Sciences, Stockholm University, Forum 100, SE-16440, Kista
{martin.henkel,js}@dsv.su.se
[5] University of Rostock, Albert-Einstein-Straße 22, Rostock, Germany
hasan.koc@uni-rostock.de

Abstract. Enterprises operate in dynamically changing environments, consequently, information systems (IS) are a subject to continuous changes. The IS need to be adjusted to different customers, markets and run-time contexts affecting their design and delivery. Variability and customization of the information systems has become a serious challenge. Capability Driven Development (CDD) has been proposed as a new approach for dealing with variability. CDD allows delivering enterprise capability in dynamically changing circumstances - contexts. To model enterprise capabilities, concepts defined in the Capability meta-model are used. This paper discusses initial Capability modeling experiences. The main emphasis is put on the Capability design phase of the CDD approach. Several empirical cases are included to share the Capability modeling experiences, to validate the Capability meta-model as well as to discuss the difficulties of Capability modeling, open issues and future work.

Keywords: Capability, variability, capability meta-model, enterprise modeling.

1 Introduction

Capability is the ability and capacity that enable an enterprise to achieve a business goal in a certain context [1]. From the Information Systems (IS) perspective the Capability allows delivering business value in a dynamically changing context. The context provides information about the present status and environment: people, places, devices, things etc. [2]. To increase the effectiveness of IS delivery and to increase user satisfaction, IS should be aware of the run time context and should be able to adapt to a certain context by delivering the service according to each user's needs. This is known as context-awareness [3]. Our stance is that application context

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 1–14, 2014.

needs to be considered in organizational and IS development, hence, a new approach for developing context-aware systems to capture context dependent variability in business designs, e.g. in terms if business processes and resources, has been proposed under the name of Capability Driven Development (CDD) [4].

The conceptual foundations of the CDD approach are defined by the Capability meta-model. The Capability meta-model combines elements of an enterprise modeling, context modeling and Capability modeling. The initial versions of the Capability meta-model have been proposed in [1, 4]. The further elaboration of the Capability meta-model and a methodology supporting the CDD approach is performed has been carried out in the FP7 project "CaaS: Capability as a Service for digital enterprises". This paper reports application of the Capability meta-model for several empirical cases. During the Capability modeling in practice, the Capability meta-model has been iteratively refined and the initial validation of Capability meta-model has been performed as part of the capability modeling at the use case companies of the CaaS project.

The objective of the paper is to evaluate the expressiveness of the Capability meta-model by applying it to different empirical cases. The paper discusses the initial Capability modeling experiences, observations, open issues and future work on Capability modeling.

The rest of the paper is structured as follows: Section 2 gives a brief overview of the Capability meta-model. In Section 3 Capability modeling of 3 example cases are presented. Section includes the description of each case, Capability model of each case and a method for Capability modeling common to all example cases. An overall observations, difficulties and open issues are provided in Section 4. In Section 5 related work on variability in IS are discussed. Section 6 summarizes the conclusions about Capability modeling experiences and gives and insight to a future work.

2 Capability Meta-Model

The Capability meta-model is designed following an integration approach using Enterprise Modeling (EM) as a starting point. The EM elements such as goals, processes and resources are used to design a business service underlying the capability. That is augmented by the capability design and delivery context to show dependence of capability delivery on the contextual factors. Making the business service contextual causes variability management problems, which is addressed by providing methods for managing and reducing variability in the form of patterns. The capability delivery performance is also context dependent, and it must be continuously monitored using appropriate measures. Therefore, the measurement elements are also included in the capability meta-model.

In summary, the key components of the Capability meta-model are the following: 1) enterprise model; 2) context; 3) patterns; and 4) measurements. The EM component starts with the representation of goals, and processes realizing the goals using required resources. These are essential components of business planning, and their relationships as presented in the figure are common to many EM approaches. The capability design is built on commonly used EM concepts, and it formulates the requirements for the ability of accomplishing the business goals in the changing circumstances. The For Enterprise Modeling (4EM) approach [5] is used here for the EM purposes.

The meta-model suggests modelling all properties of a situation as context. The meaning of the context is perceived similarly as in [6]. The context is any information that can be used to characterize the situation of the entity. It encompasses the information characterizing the situation in which a business capability should be provided. Fig. 1 shows a detailed capability meta-model. A capability is designed to be adequate for certain context situations represented by a Context Set, i.e. a range of Context Element values. Each Context Set can be materialized with a number of Context Situations according to the specification of the set by observing or measuring different individual Context Elements. The Pattern component describes an actual solution for realizing a Capability. Each pattern describes how a certain Capability is to be delivered within a certain Context Situation and what resources, process, and IS components are needed. Patterns typically describe which Process Variants should be used in accordance to a Context Set. At runtime patterns are applied according to the Context Situations representing a set of actual Context values with their Measurable Properties. The Context KPIs are used to monitor at runtime whether the pattern applied for capability realization is still valid for the current context situation. If the pattern is not valid, then capability realization should be dynamically adjusted by applying a different pattern, by reconfiguring the existing pattern (i.e., changing a utilized process, reassigning resources etc.), or by aggregating several patterns into a new pattern. Technically, the context information is captured using a context platform in a standardized format.

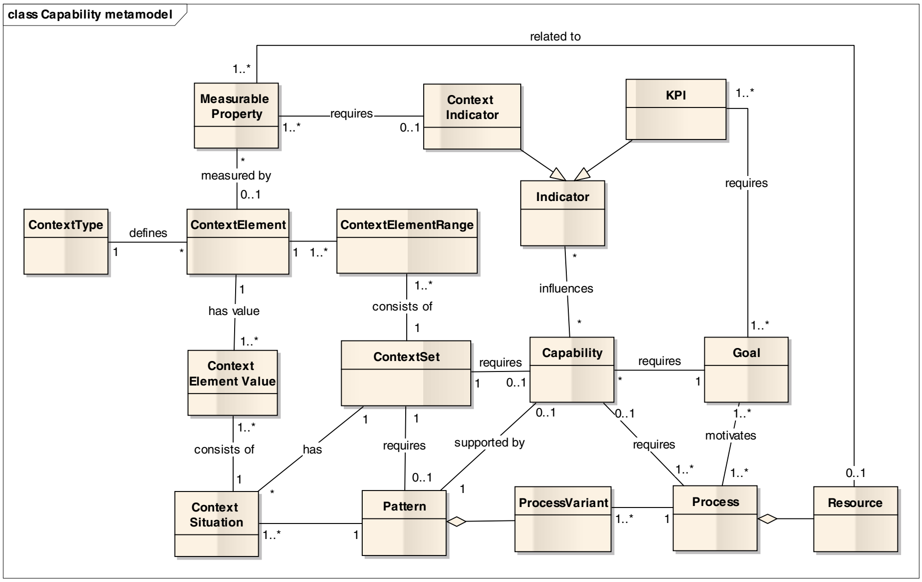


Fig. 1. Capability meta-model

Furthermore, KPIs should be set up to measure the achievements of goals. The Context KPI component defines desired KPIs that can be related to Contexts and measured using the Measured Property component, which is of a vital importance for monitoring capability delivery. In this regard, we envision that in real application cases a collection of Measurable Properties might have to be established to measure a specific Context Element.

3 Modeling Cases

In order to evaluate suitability of the Capability meta-model for representing real life capability design problems, the meta-model is instantiated to for three industrial use cases. The capability models are developed in two ways: 1) an initial capability model for the use case is developed by the academic partners and it is subsequently commented and refined by the use case company; and 2) an initial capability model for the use case is developed by the use case company following the initial guidance provided and it is subsequently jointly revised by the academic partner and the use case company. The academic partners are a part of the CaaS project team working on theoretical foundations for CDD and providing support for industrial partners in modeling capabilities and utilizing the CDD approach in their business cases.

Due to the space limitation, only the most significant parts of the capability models are represented in this paper. Each use case has its own points of emphasis corresponding to the main components of the meta-model, and the instantiations are shown for these points of emphasis.

3.1 Description of Industrial Cases

The capability models are created together with industrial partners SIV AG (SIV), Everis (EVR) and FreshTL collaborating with its business partner Danaos Management Consultants (DANAOS). In order to check different representations of the Capability Meta-model, industrial cases are chosen from diverse business domains – electricity, finance and e-governance. SIV is a business process outsourcing (BPO) company in Germany. SIV faces the challenge of continuously changing business contexts affected by new regulations, bylaws and other circumstances. This leads to a need for the context-aware solutions that deliver business value to an ever-changing market. The underlying scenario of SIV's business case is the exchange of data between two standardized market roles, the grid operator and the energy supplier. This business case is called MSCONS. Run-time adjustments may be applicable for having the system allocating decision workers from one task to another.

The EVR use case demonstrates the potential of Service Oriented Architecture (SOA) capability management in a community where various factors need to be considered and actors involved, e.g. SMEs, multinational corporations, diverse public administration's laws, regulations, administrative consortia and calendars, as well as various technological tools. The use case is based on the public sector and main emphasis is put on electronic services provided to municipalities and used by citizens and companies. EVR in the SOA platform provides a service catalogue with up to 200 services (including draft services). Approximately 100 of them are in active use in 250 municipalities in Spain. In time period from 2008 – 2013, at least 1023673 citizens of Spain are using the SOA platform. The marriage registration is one of the services. The purpose of the marriage registration submission online service capability is to provide an electronic service to citizens where citizens can apply for marriage registration in their municipality, to check the available dates and times in municipality's marriage registration institution, to find out the amount of the marriage registration tax.

DANAOS is a software and services company specializing in the maritime industry for the past 30 years. One of the business areas of DANAOS is the provision of services for regulatory compliance. The potential of applying CDD approach for DANAOS is to develop capabilities for Rule Compliance management, falling under the umbrella of "Electronic Maritime Strategic Framework" (EMSF). The EMS framework includes a number of services of great variety that DANAOS aims to provide to port authorities as well as vessels. The main focus area is the tight integration and co-operation in the fragmented field of regulatory compliance in the maritime domain. Regulations are created by numerous different bodies, with little co-operation between them. As such, there is a significant lack of cohesion between the vast array of regulations and the possibility of conflicting regulations is very real. Every shipping company must comply with the regulations of a particular port, involving a large number of documents about various aspects that are of relevance to approximately 15,000 ports world-wide. Failure to comply in a given rule may have a tremendous impact in the ship owner's ability to operate. Due to that DANAOS wishes to develop capabilities for managing regulatory rules specific to different countries and ports.

3.2 Samples Capability Models

The meta-model instantiations for the MSCONS and marriage registration services are shown Figure 2 and Figure 3, respectively. The instances shown in the models are composite instances, which are further elaborated in sub-diagrams. These sub-diagrams are expanded in the following section.

As represented in the model, the MSCONS capability's goal is to decrease process costs of MSCONS market communication. The capability is supported by the MSCONS message import process, which has multiple process variants. The capability context set and indicators are also defined. The capability is supported by a newly established pattern consisting of the process variants. The use case owners proposed that the dependence of the MSCONS capability of other capabilities also should be represented in the process model though it was decided to represent such dependences by using patterns in definition of the process variants. The EVR and DANAOS capability models at the composite level are structured similarly to the SIV's capability model.

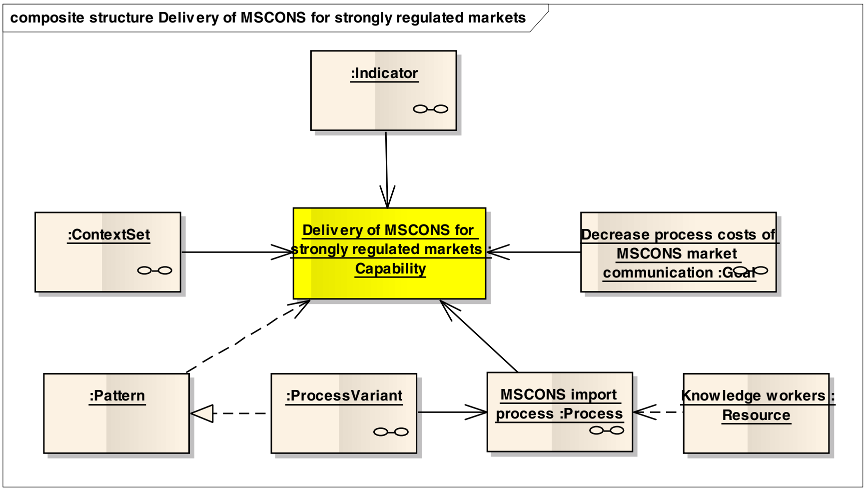


Fig. 2. The meta-model instantiations for SIV use case

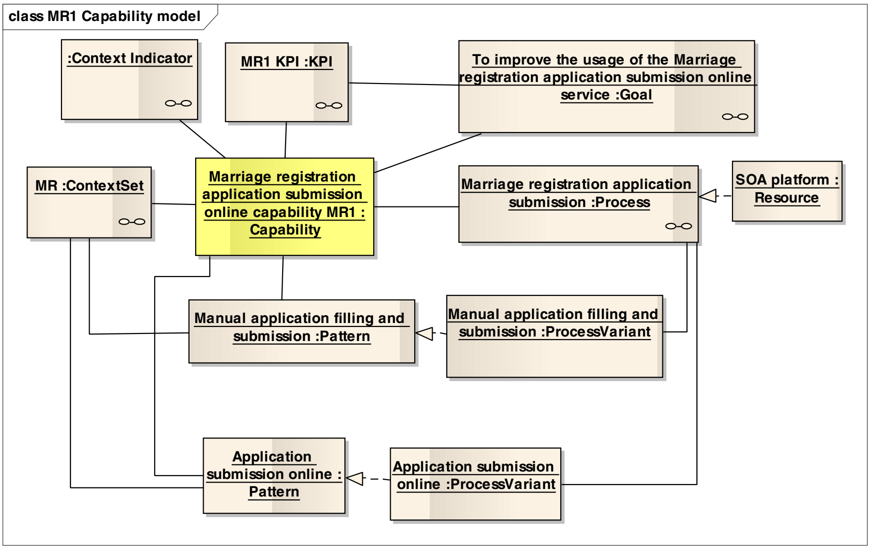


Fig. 3. The meta-model instantiations for the EVR use case

The components of the capability meta-model are instantiated in the sub-diagrams of the composite capability models. Figure 4 shows the process model for the DANAOS use case. It is important to note that the business service of compliance checking at the port is defined without indication of the context dependence. The context dependence is further elaborated in the solution part of the capability design including patterns and the process variants.

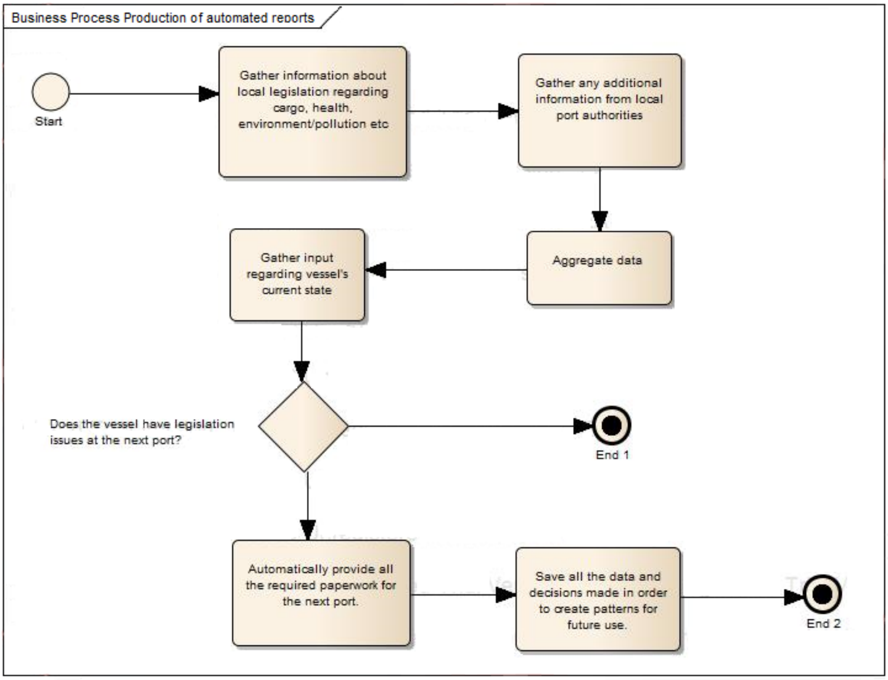


Fig. 4. The enterprise modelling component of in the DANAOS use case

In order to represent the business service context-awareness and ability to deal with changing circumstance, the context is identified. The context definition for the EVR and SIV use cases are given in Table 1 and Table 2. The context elements define factors affecting the capability delivery variability and they have specific business meaning. The measurable properties are actual observations made during the capability design and delivery, and expressions are used to transform these observations into business values as defined in the context element range.

Table 1. The context component as defined for the EVR use case

Context element	Context type	Context element range	Measurable properties
Marriage registration institution load	Dynamic	{High load, medium load, low load}	Percentage of institution load
Time of the year	Dynamic	{Winter, Spring, Summer, Autumn}	Time of the year
Type of the day	Dynamic	{Working day, weekend, national holiday}	Type of the day
Calendar of events	Dynamic	{High impact, Neutral impact, Low impact}	Types of events per calendar unit Number of events per calendar unit.

Table 2. The context component as defined for the SIV use case

Context element	Context type	Context element range	Measurable properties
Backlog size	Semi static	{High load, medium load, low load}	Defined size in customer contract
Message Type	Semi static	{MSCONS, UTILMD}	Type of the message
Regulatory Environment	Static	{EU, Non EU}	Latitude, Longitude
Message Version	Semi static	{2.0, 2.1, 2.2, 5.0}	Version of the message

The measurement component of the capability models focuses on indicators used to measure capability delivery from the business perspective as well as to monitor capability delivery circumstance. The measurements (Fig. 5 for the SIV use case) are tightly interrelated with the context and EM components. In the MSCONS capability model the measurements are used to monitor achievement of the business goals (coming from the EM component) and to monitor the context (defined in the context component).

Process variants are one of the mechanisms for representing variability (Fig. 6). SIV has to provide specific processes for different market roles, markets and types of commodities. The process variants might be needed for different combinations of these differentiators what is not represented in the current model. The decision concerning the process variant to be implemented is context dependent and the process variants are related to the context set via capability delivery pattern.

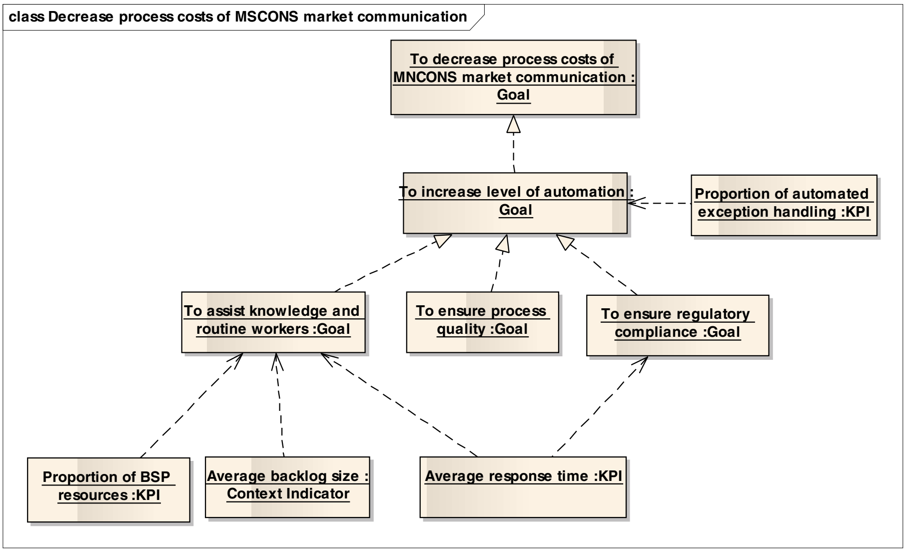


Fig. 5. The measurement component of the MSCONS capability model

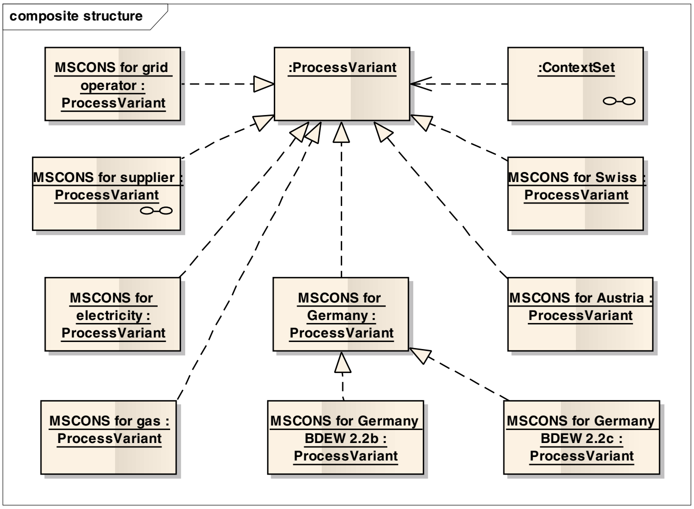


Fig. 6. Process variants of the MSCONS for strongly regulated markets capability

Fig. 7 expands the MSCONS process variant for the supplier role irrespectively of the commodity and regulatory environment. In order to be able to construct a large number of the process variants, reusability is essential. That is demonstrated by using patterns to compose the processes into the process variants. As a result existing patterns can be used in developing new process variants for supporting the capability, and once the process variants are defined they can be packaged as new patterns along with pattern attributes.

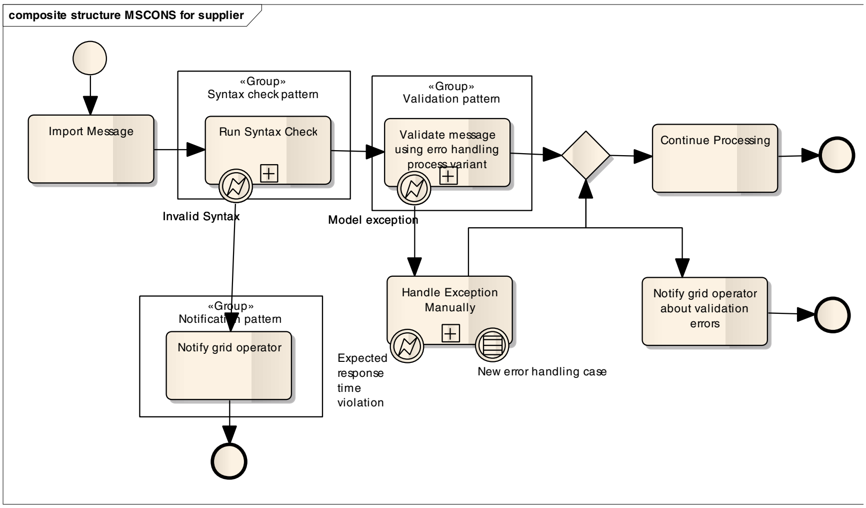


Fig. 7. MSCONS for supplier process variant with patterns

The Validate message using error handling process variant activity is a sub-process activity and this activity itself includes process variants, where each variant is intended for handling specific errors. To name an example, a customer might have a service contract with SIV that requires handling only some types of exceptions by the Business Service Provider (BSP) (for instance, messages of certain types or received from specific market roles). Moreover, the messages are assigned to BSP if the dynamic context indicator measuring the exception handling backlog size exceeds a certain threshold value.

4 Observations

We have reported initial practical experiences in developing capability models on the basis of the Capability model. The capability models are developed for all use cases, and the use case partners have approved the models as sufficiently describing their business problem. Applying the Capability meta-model to different business cases allows validating the completeness and unambiguous interpretation of the meta-model. The capability models were created in several modeling sessions with each industrial partner. The capability models, modeling observations and open challenges were discussed in the CaaS project meetings. The observations are gathered from three perspectives – academic partners involved in capability modeling for particular business cases presented in this paper, industrial partners representing their business cases and evaluating how the CDD approach satisfies their business needs and other project partners which were not directly involved in capability modeling, but which are working on theoretical and practical foundations of the CDD approach.

Several differences have been observed what should be taken into account during the further elaboration of the capability modeling and capability driven development. Majority of the meta-model elements and their associations are used to instantiate the Capability meta-model. The resource element is not used in some of the capability models. Information such as customer location, institution load and message type is captured as context elements. The attributes of the entities are measured by measurable properties and the allowed values to these elements are captured with the context element range concept. The context element and measurable property are related one-to-one or one-to-many, especially, for the context elements known at the design time. The context type is not always used in the model because its efficient use depends upon availability of taxonomy of context type.

The process variant element is used: 1) to represent variants of the capability process (SIV case); and 2) to represent variation within the process variant itself (EVR case). The former can be referred as to global process variant while the latter can be referred as to local process variant. The global process variants are represented using the Process Variant element. They are one of the means for representing variability, which can be caused by context awareness or some other factors. The local process variants can be perceived as non-standard way of business process modelling. Therefore, it is not directly represented in the capability meta-model. It is assumed that local process variants are useful if capability delivery process has a large number of different process variants. Rules for developing process variants should be defined at the design time of the capability.

The pattern element is also used to represent: 1) the solution supporting the capability; and 2) reusable components used in design of the process variants. Similarly, as with process variants the former can be referred as a global pattern supporting the capability as a whole and the latter can be referred as a local pattern supporting parts of the process. The global pattern is created for every new capability developed and latter can be used in design of new capabilities. The local patterns are retrieved from the pattern repository and are used to create process variants. The Context Situation element is not used in Capability models as it is relevant only to the Capability delivery phase and is not known during the Capability design phase. The Resource element is used only in the capability models where resources have significant impact on variability and delivery adjustment. In the SIV's capability model the knowledge workers play an important role on delivering capability while in the EVR's capability model all processes are done automatically by data base procedures thus human resource is not relevant to the EVR case.

For the MSCONS use case, capability has been divided into two different capabilities (dynamic support in strongly regulated markets/ dynamic support in weak regulated markets) where each Capability needs to satisfy different business goals and each Capability has a different process for delivering this capability. In order to operationalize the capabilities, standard processes are adjusted in accordance with the changes in the business environment. The adjustments are realized via process variants, which are applied to react to the anticipated changes in context and to adjust the Capability delivery. In the meta-model a direct relationship between context components and process variants is missing, the context set "is required by" a pattern, which in turn consists of process variants. It might be useful to introduce an association between context components and process variant or to refine the relationship between the context set and pattern.

In the EVR business case, one out of many public services provided by the SOA platform is perceived as a capability. That indicates a high level of detail of the capability. The SOA platform by itself has a capability to execute the services and each service has a capability to be executed and to deliver a certain result for the service user. This is a proof of the different levels of how capabilities can be modeled. In the EVR use case, the SOA platform includes particular service execution capabilities, which leads to a composite structure of the Capability and suggests adding a relationship to the Capability itself.

In the DANAOS case an additional element is added to the capability model – concepts. It contains concepts of a rule compliance management process as well as involved in compliance management. The DANAOS case shows that an enterprise model could be integrated with the Capability meta-model.

One of the main purposes of this paper is to evaluate the expressiveness of the Capability meta-model. In the context of conceptual modeling "the expressiveness refers to things said in a description or sayable in a description language" [7]. The expressiveness can be evaluated referring to the purpose of conceptual modeling. The purpose of the Capability meta-model is to describe the conceptual bases of the CDD and to specify concepts used in capability modeling. To evaluate whether this purpose is achieved or not, mapping has been done. The stakeholders who have knowledge about each problem domain from the described use cases declared that all concepts that were used to describe the problem domain, goals and solutions can be mapped to the concepts of the Capability meta-model. Some concepts of the meta-model can be interpreted differently; therefore more detailed methodological issues should be determined. But that does not affect the expressiveness of the meta-model in general. According to that we conclude that the Capability meta-model is enough expressive and sufficient.

5 Related Work

Variability has become a serious challenge in IS development and delivery. Variability allows managing commonalities and differences in software products and to adjust software delivery for different markets and different versions [8]. Variability explanations in the IS context include such terms as “variation point”, “process variant” and “customization” [9][10][11]. Each variation point requires a mechanism to capture the variability in order to deliver the desired software product. If the software is delivered to different organizations and markets, high level of customization becomes a serious challenge.

Model Driven Development (MDD) has been proposed as an approach dealing with variability [12]. MDD has a low abstraction level of software development artefacts and it lacks strong methodology which limits the MDD to be an appropriate approach managing variability in IS [13]. The usage of the decision models has been proposed as an approach for full software lifecycle management or variability [14]. Decision models usually require specific notations which complicates their usage in general cases. Schmid and John [15] present a variability management approach that is independent of any specific notation and allows describing variation points in general in a rather abstract level. The Propov approach [16] is another approach which supports capturing variability in business process models and provides flexible solution for business process management by inserting, deleting and moving process model fragments or modifying attributes. It provides some best practices of representing variation points in the CDD process model, but subsequently context dependence should be taken into account in both – system design and run time.

The CDD approach includes context situation analyzes during the software run-time phase and software adjustment to the run-time context. That means that variability does not need to be fully designed during the software design phase or managed with system configuration, which is an advantage of CDD approach as an approach dealing with variability. Customization can be done automatically as well by monitoring the static and dynamic context for each organization or market and adjusting software to each organizations requirements, legislation and other issues.

The CDD approach has been proposed as an approach dealing with variability, but it combines a number of advantages. In the CDD approach capabilities are linked to IS models. Enterprise modeling is widely used in IT industry for many years to represent the business in a form of formal models. Modeling bridges the communication gap between a business and its IT systems [17], but there are lack of methods for automated transformation from models to executable IS [4].

Context represents the business service and business capability context-awareness and ability to deal with changing circumstances. Context data usage in IS development has been discussed for many years, mainly focusing on context capturing and adaptation. But context studies lack reasoning about predicting and reusing context and context aware elements during the system run-time phase [18]. The CDD approach includes the identification, prediction and reuse of the context data.

6 Conclusion

In this paper we use the Capability meta-model to develop the Capability models for three industrial cases in order to discuss capability modeling experiences and evaluate expressiveness and sufficiency of the Capability meta-model. We conclude that Capability meta-model is sufficient for modeling different business cases. It flexible enough to represent different business cases and to adjust to the varying needs of the industrial partners. Though, several difficulties and clarifications required are observed during the capability modeling. One of the main open issues is related to deciding on appropriate level of decomposition the capability concept should be defined.

Future work on capability modeling has two main directions: 1) detailed methodological support should be provided in order to ensure unambiguous application of the capability meta-model; and 2) providing information for semi-automated generation of context-aware components to be integrated with capability delivery applications. Concerning the former issue, our initial capability modeling experiences show that different modeling pathways are used for different used cases what might affect the final result. Concerning the latter issue, a trade-off between the level of detail required to define capabilities from the business view and the level of detail required to support capability implementation should be identified.

References

1. Zdravkovic, J., Stirna, J., Henkel, M., Grabis, J.: Modeling Business Capabilities and Context Dependent Delivery by Cloud Services. In: Salinesi, C., Norrie, M.C., Pastor, Ó. (eds.) CAiSE 2013. LNCS, vol. 7908, pp. 369–383. Springer, Heidelberg (2013)
2. Korpipaa, P., Mantyjarvi, J., Kela, J., Keranen, H., Malm, E.J.: Managing context information in mobile devices. IEEE Pervasive Computing 2(3), 42–51 (2003)
3. Chihani, B., Bertin, E., Jeanne, F., Crepsi, N.: HEP: context-aware communication system. International Journal of New Computer Architectures and their Applications (IJNCAA 2011) 1(1), 15–24 (2011)

4. Stirna, J., Grabis, J., Henkel, M., Zdravkovic, J.: Capability Driven Development - an Approach to Support Evolving Organizations. In: Sandkuhl, K., Seigerroth, U., Stirna, J. (eds.) PoEM 2012. Lecture Notes in Business Information Processing, vol. 134, pp. 117–131. Springer, Heidelberg (2012)
5. Sandkuhl, K., Wißotzki, M., Stirna, J.: Unternehmensmodellierung: Grunlagen, Methode und Praktiken (German Edition). Springer, Heidelberg (2013)
6. Dey, A.K.: Understanding and using context. Personal and Ubiquitous Computing 5(1), 4–7 (2001)
7. Patig, S.: Measuring Expressiveness in Conceptual Modeling. In: Persson, A., Stirna, J. (eds.) CAiSE 2004. LNCS, vol. 3084, pp. 127–141. Springer, Heidelberg (2004)
8. Galter, M., Avgeriou, P.: The Notion of Variability in Software Architecture – Results from a Preliminary Exploratory Study. In: Proceedings of the 5th Workshop on Variability Modeling of Software-Intensive Systems, pp. 59–67 (2011)
9. La Rosa, M., Aalst, W.M.P., Dumas, M., Milani, F.P.: Business Process Variability Modeling: A Survey, QUT E-prints, `http://eprints.qut.edu.au/61842/1/BPM-13-16.pdf` (accessed March 8, 2014)
10. Hallerbach, A., Thomas, B., Reichert, M.: Capturing Variability in Business Process Models: The Provop Approach. Journal of Software Maintenance and Evolution: Research and Practice 22(6-7), 519–546 (2010)
11. Bosch, J., Florijn, G., Greefhorst, D., Kuusela, J., Obbink, J.H., Pohl, K.: Variability Issues in Software Product Lines. In: van der Linden, F.J. (ed.) PFE 2002. LNCS, vol. 2290, pp. 13–21. Springer, Heidelberg (2002)
12. Santos, L.A., Koskimies, K., Lopes, A.: A Model-Driven Approach to Variability Management in Product Line Engineering. Nordic Journal of Computing 13(3), 196–213 (2006)
13. Henkel, M., Stirna, J.: Pondering on the Key Functionality of Model Driven Development Tools: the Case of Mendix. In: Forbrig, P., Günther, H. (eds.) BIR 2010. LNBIP, vol. 64, pp. 146–160. Springer, Heidelberg (2010)
14. Atkinson, C., Bayer, J., Bunse, C., Kamsties, E., Laitenberger, O., Laqua, R., Muthig, D., Paech, B., Wust, J., Zettel, J.: Component-based Product Line Engineering with UML. Component Software Series. Addison-Wesley (2001)
15. Schmid, K., John, I.: A customizable approach to full lifecycle variability management. Science of Computer Programming 53(3), 259–284 (2004)
16. Hallerbach, A., Bauer, T., Reichert, M.: Capturing variability in business process models: A Propov approach. Journal of Software Maintenance and Evolution: Research and practice 22, 519–546 (2010)
17. Zhao, L., Letsholo, K., Chioasca, E.V., Sampaio, S., Sampaio, P.: Can business process modeling bridge the gap between business and information systems? In: Proceedings of the 27th Annual ACM Symposium on Applied Computing, Trento, Italy (2012)
18. Vale, S., Hammoudi, S.: COMODE: A framework for the development of context aware applications in the context of MDE. In: Proceedings of the 2009 4th International Conference on Internet and Web Applications and Services, pp. 261–266 (2009)

IT Outsourcing Relationships in Swedish Public Organizations

Ilir Çelaj[1] and Lazar Rusu[2]

[1] School of ICT, Royal Institute of Technology (KTH), Stockholm, Sweden
[2] Department of Computer and Systems Sciences, Stockholm University, Sweden
celaj@kth.se, lrusu@dsv.su.se

Abstract. Information technology (IT) is not only beneficial for the private sector but has a lot to offer for the public sector as well. However, most of the studies in IT outsourcing relationships consider only the private sector and do not pay attention to the public sector. In this context, the research has looked to determine the influential factors in IT outsourcing relationships in Swedish public organizations like are Swedish municipalities and classified them according to the importance. The research methodology used has included a survey among IT outsourcing decision makers from Swedish municipalities and a case study in a Swedish municipality. The survey has revealed that trust, cooperation, commitment, communication, and flexibility are the top five most important factors which are influencing the provider-recipient IT outsourcing relationship in Swedish municipalities. While the case study has revealed that all the factors that were identified during the literature review are influencing the provider-recipient IT outsourcing relationship in Swedish municipalities plus the identification of two new factors which are legislation and economy/money.

Keywords: IT outsourcing, IT outsourcing relationship, IT outsourcing decision makers, service recipient, public sector, Swedish municipalities.

1 Introduction

Information Technology (IT) outsourcing has been around for quite some time and will continue to be in the near future. IT outsourcing is known as "*a joint decision to sign a contract which stipulates that the IT-supplier should perform information systems development activities for the client over an agreed time period, irrespective of where the IT supplier is located*" [6, p. 463]. IT outsourcing was firstly used by some IT managers in 1970s in order to deal with increasing costs [41]. The research on the domain of IT outsourcing has been done for several years till now. During the 80's and 90's, there was a lot of scientific research on outsourcing [57]. Due to the global social changes, outsourcing has become an interesting issue and has raised much interest in the last 20 years [55]. We are all witnesses that IT outsourcing is becoming very important and that organizations in both public and private sector are outsourcing some of their goods and/or services [59]. In IT outsourcing, the external vendors which are providing services, system

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 15–29, 2014.

functions, software, and/or hardware are known as service providers while companies or organizations that are contracting them are known as service recipients [7]. As we could see, there are many IT outsourcing drivers and challenges. One of the challenges in IT outsourcing that it is very important and that needs to be considered by the IT outsourcing decision makers is the need to work with providers and to effectively manage the relationship with them. According to [22] it is vital for IT outsourcing to understand the IT outsourcing relationship. Although it is considered important, IT outsourcing relationship has received least attention in the IT outsourcing domain [22] and the reasons why in the 1990s IT outsourcing relationship did not receive so much attention remain unclear and unexplainable. However, the situation has changed in the recent years as an increased number of research studies in IT outsourcing relationship regarding the provider-recipient relationship have been reported [4]. However, most of them consider only the private sector and do not pay attention to the public sector. Moreover according to [52] the studies in these sectors should be approaches differently as they are not the same and they have different results. According to [38] and [52] IT outsourcing relationship management often determines IT outsourcing success. Therefore, IT outsourcing relationship management has become an important research topic to be studied in IT outsourcing area. However, as we have noticed public sector organizations are struggling in IT outsourcing relationship management because there are much less research studies done in this sector than in the private sector [52]. As we can see, the problem is that there is not so much research done previously in the public sector about the IT outsourcing relationship. Furthermore we have not found in the research literature any previous studies which consider the factors that have an influence on the provider-recipient IT outsourcing relationship in Swedish municipalities. The decision on which factors should be given more importance maintains a challenge since there is no classification of factors according to the importance. In order to address this problem, this study will look to find an answer to the following research question: *Which are the factors that influence the IT outsourcing relationship in Swedish municipalities and which of them are the most important?* The intent of this study is to investigate the IT outsourcing relationship from the service recipient's perspective in Swedish municipalities as a specific type of public sector organizations. Furthermore, identify and classify factors that are deemed as important and that influence the provider-recipient IT outsourcing relationship.

2 Research Background

2.1 IT Outsourcing in Public Sector Organizations

There are many definitions for IT outsourcing. In this study, the authors will try to summarize and mention some of the definitions made by several researchers. For example IT outsourcing is defined as the process of transferring the business function to an external provider or organization that is specialized in providing the particular function is known as outsourcing [17]. Another definition of IT outsourcing is concerning the decision to have some information systems activities and functions provided by an outside supplier and in exchange for money, the supplier provides these services for a period of time while the activities and functions that can be outsourced include network systems, operations, and so on [25, 34, 29, 22, and 45]. The definition that we have agreed to use in this study is defining IT outsourcing as "a joint decision to sign a contract which stipulates that the IT- supplier should perform information systems development (ISD) activities for the client over an agreed time period, irrespective of where the IT supplier is located" [6, p. 463]. The private and public sector organizations have some common values that they share. These values include accountability, responsiveness, incorruptibility, and honesty. However, private and public sector organizations have their differences as well. The values of the public sector organizations do not have to do that much with money, while the goals of private sector organizations are strongly connected to sustainability and profitability [13]. Public sector organizations have the ability to adapt to the requirements of the internal or external environment, despite the fact that bureaucracy and formalization have been part of the culture of the public sector organizations for a long time. The fact that public sector organizations have started to get more and more involved in IT outsourcing means that they are able to adapt. This ability to adapt tells us that the behavior of the public sector organizations has some specifics that may be worth considering and that studying the IT outsourcing in public sector organizations is necessary [13]. The competition and pressure increases in the private sector organizations whenever there is a shortage of resources. However, this is not the case with public sector organizations. Public sector organizations increase the collaboration and share the resources between them. The fact that public sector organizations do not compete with each other makes it easier for them to collaborate with each other. This collaboration is hard to be achieved in the private sector organizations [13]. Service providers in public sector organizations are chosen based on a process that is very strict and that includes tendering. The public sector organizations must always sign a contract with the service provider that offers the lowest price which is not the case with private sector organizations. Employing family members is normal in private sector organizations while employing them in public sector organizations is considered as non-ethical [13].

2.2 IT Outsourcing Relationship Factors

Several factors have been identified during the literature review as factors that are important and influence the IT outsourcing relationship. In total a number of 30 factors have been identified in the literature review in IT outsourcing relationship. However, only 17 have been considered since the other 13 factors are not applicable in both sectors and have been mentioned only in one article. Since this study considered research literature that had to do with both private and public sector, it was important to indicate which authors carried out the research in the public sector. Besides [42] which stated that the research was carried out in the public sector, all other authors either carried out the research in the private sector or just did a literature review and the sector could not be implied. As a result of the literature review we have considered 17 influential factors which are followings: cooperation, trust, culture, personal and social bonds, communication, conflict, commitment, coordination, dependency, flexibility, age of the relationship, knowledge/information sharing, top management support, participation, consensus, performance management, and contract management. The factors which have been only once mentioned in the research literature and which have been excluded are: satisfaction, outsourcing strategy, due diligence, contract development, governance, integration, participation, joint action, shared vision, attraction, expectations, power, attitudes and assumptions.

A description of the influential factors in IT outsourcing relationship considered in this study is included below.

Cooperation in a relationship means that both the client and the vendor work together in order to achieve their interests in the end [26]. Cooperation and coordination are similar in concept but the difference is that cooperation captures the nature of the relationship in a higher level of abstraction [14]. Cooperation is very important for the IT outsourcing relationship [23, 3, 5, 42, 1, 22, and 16].

Trust means that an organization has faith in the partner organization that it will behave properly and will not try to take advantage of the relationship [26]. Trust is significant for the provider-recipient relationship [42, 24, 1, and 23] and has a big effect on its success and the overall IT outsourcing success [14, 31, and 22].

Culture is a collection of norms, values, and behavior patterns [8]. Cultural understanding should be present in outsourcing relationships [58] because culture is considered important for the provider-recipient IT outsourcing relationship [21, 31, 22, 23, and 14].

Personal and social bonds have been defined as personal confidence, friendship, and familiarity built through exchange [47]. It is significant for the provider-recipient relationship [22, 23, and 50].

Communication has been defined as the sharing of informal and formal information [51]. It is considered as an important aspect of the provider – recipient IT outsourcing relationship [14, 48, 42, 1, 9, 10, 22, 3, 31, 21, and 23].

Conflict includes disagreements over execution of the contract, procedures, roles, and goals [10]. Conflict is one of the IT outsourcing relationship factors [22, 23, 14, 31, 3, and 21] and it influences the success of the working relationship [1, 42].

Commitment can be defined as the will to remain in the relationship [49]. Commitment is an attribute that influences the success of the working relationship [1, 42] and it is crucial for the provider-recipient IT outsourcing relationship [22, 23, 14, 10, 21, 20, and 31].

Coordination has been defined as the accomplishment of the tasks which parties have agreed upon with each- other through management of mutual dependencies [54, 36]. It is important for the provider-recipient IT outsourcing relationship [43, 14, 40, and 31].

Dependency is the dependence of each of the parties participating in the relationship on the other party for the occurrence of their goals, acts, and behaviors [43]. It is acknowledged as a factor that influences the provider- recipient IT outsourcing relationship [42, 1, 3, 20, 22, 31, 23, 21, 14, and 10].

Flexibility is defined as willingness of the both parties engaged in the relationship to make changes in the case when circumstances have changed [19]. Flexibility is is important [11, 35] because of the unpredictable environment and because of the length of outsourcing contracts [14].

Age of the relationship is a factor which influences the IT outsourcing relationship [31, 2, 56, and 53]. It is more likely for a long-term relationship to continue than a short-term relationship if the participants have developed it over time [12].

Knowledge/information sharing can be defined as the amount of information which is communicated between the parties [32]. Knowledge/information sharing is a factor which influences the IT outsourcing relationship [1, 20, 31, and 33]. Competitive advantage can be created whenever there is strategic sharing of information which is important for both service receiver and service provider organizations [28].

Top management support can be defined as the degree of understanding and support provided by executives of both parties for the provider-recipient IT outsourcing relationship [12]. Top management support is a factor which influences the IT outsourcing relationship [44, 31] as is important for the establishment of trust and for the exchange of knowledge between the parties [1].

Participation has been defined as the level of involvement of the provider and the recipient in the outsourcing relationship [12]. It is one of the factors which influence the provider-recipient IT outsourcing relationship [18, 31, and 16] as it makes it possible for the relationship to be sustained over time [20].

Consensus has been defined as level of agreement between service recipient and service provider [39]. Consensus is a factor that is important for the provider-recipient IT outsourcing relationship [46, 14] as it is often the case that strategic decisions made by the parties participating in the relationship reflect the mutual agreement (consensus) between them [21].

Performance management is defined as continuous measurement and monitoring of the performance of the service provider [15]. Performance management is important to make sure that supplier is fulfilling its obligations [27] and it is the main reason why clients decide to outsource support center [1].

Contract management is defined as continuous monitoring and reporting of the outsourced and delivered services against the defined criteria [15]. Contract management is an important factor for the provider-recipient IT outsourcing relationship [1, 11, 37, 27] as poor management of the contract leads to dissatisfaction and conflict in the provider-recipient IT outsourcing relationship [30].

3 Research Method

The research done in this study comprises three steps. These steps include a literature review in IT outsourcing relationship, empirical research, analysis, and conclusions. In this research we have used both a quantitative and qualitative approach. The first step includes a literature review and its purpose was to identify relevant information regarding IT outsourcing and IT outsourcing relationship that would create the basis for the research and help us to answer the research question. Then in the second step we have collected the data through a survey and an interview while data analysis and conclusions was done in the third step.

The results of the literature review were used to create a questionnaire and the interview questions. The search of the relevant research was done through well-known databases such as Science Direct, IEEE Xplore, Emerald, and AMC Digital Library as well as searching on Google scholar by using open queries and keywords like IT outsourcing, outsourcing relationship, service recipient, service provider, public sector, and municipalities. Based on the literature review, a number of 30 IT outsourcing relationship factors were identified. After performing the literature review, we have performed an empirical research which was done by using a survey and a case study. The survey has been chosen to classify the factors which were identified during the literature review according to the importance while the case study research has been chosen in order to complement the survey. We have used as a data collection techniques the questionnaires and interviews. The questionnaire is included 10 closed questions (with several alternative answers) with predefined answers in order to get precise answers. The interview has been chosen as a data collection technique in order to get detailed information from the respondent about influential factors in IT outsourcing relationship. The survey was performed between March to April 2012 by using a questionnaire that was created on the Survey Monkey website: https://www.surveymonkey.com/ followed in June 2012 by an interview with an IT manager involved in IT outsourcing decisions from a Swedish municipality.

4 Results and Analysis

4.1 Survey Results and Analysis

The data was collected through a survey done from March 28, 2012 to April 28, 2012. For this purpose a questionnaire was distributed to responsible persons for IT Outsourcing from 290 Swedish municipalities. The result of this survey has included 32 IT outsourcing decision makers like Chief Information Officer, Head of IT, IT manager, IT chief, IT strategy officer at Executive office, Group leader of the system group, ICT strategist that have replied during the period above mentioned and performed the survey. The questions which were related to the factors that have an influence on IT outsourcing relationship has been answered only by 20 or 62.5% of the total number of respondents (32) and has been skipped by 12 or 37.5% of the total number of respondents (32). The respondents had the possibility to choose between classifying the 5, 10, or 17 factors according to the importance by assigning numbers (1-17) to them as well as to add comments in the “Comments” box. But only 17 respondents or 85% have classified all the factors (17), none or 0% have classified only 10 factors, and 3 respondents or 15% have classified only 5 factors. The lowest number (1) means that a factor is more important and influences the IT outsourcing relationship the most, while the highest number (17) means that a factor is less important and influences the IT outsourcing relationship the least. Each of the numbers given from the respondents for each specific IT outsourcing relationship factor have been added and then divided by the number of respondents that have answered the question (20). In this way we have made an average for each IT outsourcing relationship factor and classify the factors according to the importance by starting with the factor with the lowest average and ending with the factor with the highest average. The classification of the factors according to the importance given by IT outsourcing decision makers from the survey done in Swedish municipalities is shown in Table 1.

Table 1. Classification of the factors in ITO relationship according to the importance

Nr.	IT Outsourcing Relationship Factors	Average
1	Trust	3.80
2	Cooperation	4.83
3	Commitment	5.42
4	Communication	6.11
5	Flexibility	6.24
6	Performance management	6.58
7	Knowledge/information sharing	7.22
8	Coordination	8.22
9	Participation	8.65
10	Contract management	9.29
11	Top management support	10.28
12	Culture	11.18
13	Dependency	11.53
14	Age of relationship	11.56
15	Consensus	12.33
16	Conflict	12.76
17	Personal and social bonds	12.83

As we can see from the Table 1 above, trust, cooperation, commitment, communication, and flexibility seems to be the most important factors that are influencing the IT outsourcing relationship in Swedish municipalities.

4.2 Case Study Results and Analysis

Case Study: In this research we have selected as case study Municipality X from Stockholm County. Due to confidentiality we cannot mention the name of the Swedish municipality neither the name of the interviewee. The interview was performed on June 25, 2012 with an IT manager of this municipality that was responsible for both IT and telecommunications, has an experience in IT of about 14 years and was involved in IT outsourcing decisions in this municipality. The municipality X has signed three main contracts for IT outsourcing with three different service providers. For taking decisions regarding IT outsourcing the municipality X has established a steering group. As we mentioned before the municipality X has three main service providers which were chosen through the procurement process that is based on the Swedish law. The biggest problem in the municipality X is that these service providers write down perfect bids and deliver completely something else. Usually, the service providers deliver only 50 to 80% of what was written in the bid.

IT Outsourcing Relationship Factors in a Swedish municipality

A presentation of the results coming out from the interview performed with an IT manager from a Swedish municipality concerning the IT outsourcing relationship factors are below included.

Cooperation: The interviewee has acknowledged the positive role of cooperation in the relationship and said that cooperation tends to be higher when there are large projects in place. Cooperation has been included in the top 5 most important factors influencing the IT outsourcing relationship in the survey research as well.

Trust: The empirical research showed us that trust is an important factor for the IT outsourcing relationship in the case that was studied. The interviewee said that "high levels of trust have a positive influence on the relationship". It was also ranked as the most important factor (on average) from 20 IT managers of the Swedish municipalities that participated in the survey.

***Culture*:** The results gathered during the case study indicate that culture as a factor is present and influences the relationship, although it is not ranked as very important in the survey research. The interviewee said that culture has a positive influence on the relationship and this is in contrast with what we are used to read. We are used to read that culture has a negative influence on the relationship because of the differences between two organizations. However, regardless of whether the type of influence, we have managed to see that culture has an influence on the IT outsourcing relationship in the municipality where the case study was performed.

***Personal and social bonds*:** The personal and social bonds have been mentioned several times in the previous studies and have been deemed as very important for the IT outsourcing relationship. During the case study, personal and social bonds have been confirmed as a factor influencing the IT outsourcing relationship in Swedish municipalities. Personal and social bonds positively influence the relationship and have been deemed by the interviewee as more important than the culture which in a way contradicts the results of the survey where 20 IT managers have participated.

Communication: According to the interviewee, communication comes as a result of cooperation. He said that "the more cooperation you have, the more communication you need". He also implied that the influence is positive as communication improves the relationship. The interviewee gave importance to the difference between formal and informal communication and made that clear several times during the interview. The case study results complemented the results which authors gained through the survey where communication was ranked among the top 5 most important factors influencing the IT outsourcing relationship in Swedish municipalities by 20 IT managers.

Conflict has been acknowledged as a factor having an influence on the IT outsourcing relationship in Swedish municipalities during the case study research. However, the interviewee had some doubts whether the influence was positive or negative. He said that "disagreements are negative at first, but they can be positive in the long run if an agreement is achieved in the end. Otherwise, if there is no agreement in the end, conflict is very negative". By analyzing this sentence, one can see that the occurrence of conflict has a negative influence, while the process of conflict resolution has a positive influence on the relationship. The author is however not interested in the process of conflict resolution but on the phenomenon of conflict occurrence itself. Therefore, we can say that conflict occurrence negatively influences the provider-recipient IT outsourcing relationship in Swedish municipalities.

Commitment: According to the interviewee, commitment is very important for the relationship. He said that “the municipality and the service provider have signed agreements and that represents the base in their commitment”. He argued that commitment influences the relationship with the service provider in a positive way. Commitment has also been ranked as one of the top 5 most important factors in Swedish municipalities by 20 IT managers that participated in the study. Having these results in mind, the author can say that commitment is not only important for the relationship in private sector organizations but for municipalities as public sector organizations as well.

Coordination is according to the interviewee, “to do with everyday business and builds up the commitment”. He also stated that coordination positively influences the provider-recipient IT outsourcing relationship. Besides this, coordination has also been ranked very high in the survey which was carried out before the case study and where 20 IT managers from different Swedish municipalities have participated.

Dependency: In the municipality where the case study was carried out, not so much attention is given to dependency. The IT manager did not talk much about the matter, although he said that “the more dependent that the municipality is on the services delivered, the more flexibility it needs”. He also stated that dependency has a positive influence on the relationship and that he would rather have one provider per service than more providers. It is really strange to see dependency positively influencing the provider-recipient IT outsourcing relationship but we cannot be the judge of that since we do not know what the IT manager had in mind. Usually dependency has been deemed negative since you depend on the service provider in order to achieve some/all of your goals. If the service provider does not deliver, then you do not achieve your goals and that has a direct negative influence. However, what it can be said is that dependency has an influence on the relationship.

Flexibility**:** It has been identified in the literature as a factor which is important for the provider-recipient IT outsourcing relationship. The municipality where the case study was carried out acknowledges flexibility as a factor which is important for the relationship. The IT manager said that “necessary adoptions require intensive communication, meetings, and discussions” and that the law is the only problem that can define how much flexible the parties involved in the relationship can be. He also stated that flexibility has a positive influence on the relationship. Flexibility has also been included in the top 5 most important factors influencing the IT outsourcing relationship from the IT managers of Swedish municipalities which participated in the survey.

Age of the relationship**:** The IT manager of the municipality where the case study was carried out stated that “if the municipality had a contract with a provider previously it is normal that it would influence the current outsourcing relationship”. He also stated that the age of the relationship has a positive influence. Although he stated that the influence is positive, the author would be a bit skeptical because it depends on the length of the contract to determine the type of influence. Whether the municipality had a contract with a specific service provider before and whether the current contract is long or short term determines the type of influence. However, the authors are going to conclude that the influence is positive, as the IT manager underlined. Regardless of the type of influence, the thing we know for sure is that there is an influence.

***Knowledge/information sharing*:** The IT manager acknowledged that the way they share information/knowledge has an impact on the relationship. He stated that the "information in the beginning is formal and has to do with how to read the contract and how to understand the services delivered". He stated that information/knowledge which is shared in later stages tends to be less formal and there is a larger focus on the services delivered. The IT manager stated that information/knowledge sharing has a positive influence on the relationship.

***Top management support*:** The IT manager has agreed on the importance of the top management support and said that "the involvement of the top management makes the relationship more formal". He said that top management was involved several times and in most of the cases problems with the service provider were present. He stated that the influence of the top management support on the IT outsourcing relationship is positive.

***Participation*:** The IT manager mentioned that "the level of involvement of the service provider and service recipient in the outsourcing relationship is positive for the relationship". However he did not however explain anything in details. We assume that this could imply that frequent involvement has a positive influence on the relationship.

***Consensus*:** The IT manager pointed out that consensus has a positive influence on the relationship and he mentioned that "the level of agreement is important and has a positive influence on the relationship. Usually, the municipality has service level agreements that determine which things should be done and in what time period".

Performance management: The IT manager did not really explain on how they perform the performance management process on their part and who is involved. However, he acknowledged its importance and said that performance management has an important influence on provider-recipient IT outsourcing relationship and that "the service provider keeps track of things and informs the municipality of the progress made".

Contract management has been identified in the literature as one the factors which are important for the IT outsourcing relationship. The IT manager of the municipality noticed that "more than half of the time is spent with the service provider in order to agree on how to monitor the contact". He stated that the communication is quite formal with reports, protocols, and discussions about how to read the contract and about how much the service provider has fulfilled it. He clearly stated that contract management has a positive influence on the relationship but did not explain how the contract management process is performed.

As we have noticed all the factors that were identified during the literature review and that influence the IT outsourcing relationship in the private sector organizations are influencing the IT outsourcing relationship in the Swedish municipalities as well.

But this study has also revealed that besides the factors that are present in the research literature, there are also other factors that are important and that have an influence on the IT outsourcing relationship in Swedish municipalities. The new factors mentioned by the interviewee in our case study include legislation and economy/money. The legislation seems to be a factor that it is influencing the relationship when it comes to the contract and negotiation part. It also seems to have a positive influence, although the influence can be negative as well depending on how it is handled. The economy/money factor seems to have an influence on the relationship as well. In fact the service delivery is influenced in a negative way whenever the provider tries to lower the expenses. Furthermore, the inability of the municipality to pay the provider also influences the relationship in a negative way. A graphical representation of the findings can be seen in Figure 1.

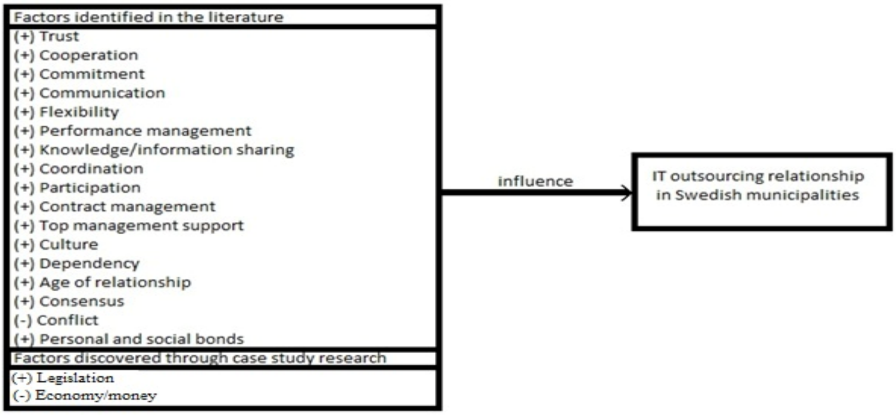


Fig. 1. Factors influencing the IT outsourcing relationship in Swedish municipalities

As we can see in Figure 1 the factors that influence the IT outsourcing relationship positively have a plus (+) in front of them, while the factors that influence negatively the IT outsourcing relationship have a minus (-) in front of them.

5 Conclusions

This study has investigated the provider-recipient IT outsourcing relationship from the service recipient's perspective in Swedish municipalities as a specific type of public sector organizations. The main purpose of this research was to identify and classify the factors influencing the IT outsourcing relationship and therefore answer to our research question. In order to answer the research question, the authors have done a survey among Swedish municipalities and also an interview with an IT manager in charge with IT outsourcing in a Swedish municipality. Additionally, the authors have provided a classification of the factors according to the importance based on the opinion coming out from the survey among IT outsourcing decision makers from municipalities in Sweden. For this purpose we have preliminary performed a literature review in order to identify the influential factors and then prepare the questionnaire and the interview questions. There are some limitations in this study therefore the results of this research would be more complete if the service provider side would be considered and not just the service recipient side. Moreover this study did not consider the relationship between different IT outsourcing relationship factors and therefore considering it would be also a good addition to the work done. The classification of the IT outsourcing relationship factors was maybe rather subjective as different people have different opinions and different definitions for the factors, so a further work could be done in that part. Moreover extending the list of the most important factors could also be done by including the factors identified during the case study (legislation and economy/money) in the classification and performing another survey. The results revealed from this research have identified the factors that are influencing the IT outsourcing relationship in Swedish municipalities which are trust, cooperation, commitment, communication, flexibility, performance management, knowledge/information sharing, coordination, participation, contract management, top management support, culture, dependency, age of relationship, consensus, conflict, personal and social bonds, and also legislation and economy/money as two new additionally factors. Moreover this research has also revealed the most important factors that are influencing the IT outsourcing relationship in Swedish municipalities which are trust, cooperation, commitment, communication, and flexibility including a classification of all the influential factors according to the importance of them for IT outsourcing decision makers.

References

1. Alborz, S., Seddon, P.B., Scheepers, R.: A Model for Studying IT Outsourcing Relationships. In: 7th Pacific Asia Conference on Information Systems, Adelaide, pp. 1297–1313 (2003)
2. Anderson, E., Weitz, B.: Determinants of continuity in conventional channel dyads. Marketing Science 8(4), 310–323 (1989)
3. Anderson, E., Narus, J.: A model of distributor firm and manufacturer firm working partnerships. Journal of Marketing 54(1), 42–58 (1990)
4. Arshad, N.H., May Lin, Y.A., Mohamed, A., Udin, A.B.: Service delivery in ICT outsourcing: relationship intensity between vendor and client. Public Sector ICT Management Review 2(2) (2008)
5. Axelrod, R.M.: The Evolution of Cooperation. Basic Books, New York (1984)
6. Bergkvist, L., Fredriksson, O.: Outsourcing terms – A literature review from an ISD perspective. In: 16th European Conference on Information Systems, pp. 458–469. ECIS, Galway (2008)
7. Beulen, E., Ribbers, P.: Control in outsourcing relationships: governance in action. In: Proceedings of the 40th HawaiiInternational Conference on System Sciences, p. 236. IEEE, Hawaii (2007)
8. Blumenberg, S., Beimborn, D., Koenig, W.: Determinants of IT Outsourcing Relationships: A Conceptual Model. In: Proceedings of the 41th Annual Hawaii International Conference on System Sciences, Hawaii (2008)

9. Cullen, S., Willcocks, L.: Intellegent IT Outsourcing: Eight Building Blocks to Success. Butterworth-Heinemann, Oxford (2003)
10. Dwyer, F.R., Schurr, P.H., Oh, S.: Developing Buyer-Seller Relationships. Journal of Marketing 51, 11–27 (1987)
11. Fitzgerald, G., Willcocks, L.: Contracts and Partnerships in the Outsourcing of IT. In: 15th International Conference on Information Systems, Vancouver, pp. 91–98 (1994)
12. Fleming, R., Low, G.: Information System Outsourcing Relationship Model. Australasian Journal of InformationSystems 14(2), 95–112 (2007)
13. Gantman, S.: IT Outsourcing in the Public Sector: A Literature Analysis. Journal of Global Information Technology Management 14(2), 48–83 (2011)
14. Goles, T., Chin, W.W.: Information Systems Outsourcing Relationship Factors: Detailed Conceptualization and Initial Evidence. In: The DATA BASE for Advances in Information Systems, pp. 47–67 (2005)
15. Gong, H., Tate, M., Alborz, S.: Managing the Outsourcing Marriage to Achieve Success. In: Proceedings of the 40th Hawaii International Conference on System Sciences, p. 239c. IEEE, Hawaii (2007)
16. Gottschalk, P., Solli-Saether, H.: Managing Successful IT Outsourcing Relationships. IRM Press, Hershey (2006)
17. Gramatikov, M.: Outsourcing of Public Information Systems, University of Sofia "St. Kliment Ohridski", `http://www.inst-informatica.pt/servicos/informacao-e-documentacao/biblioteca-digital/gestao-de-si-ti-1/aquisicao-de-bens-e-servicos/UNPAN003861.pdf` (accessed on January 17, 2013) (retrived)
18. Han, H.S., Lee, J.N., Seo, Y.W.: Analyzing the impact of a firm's capability on outsourcing success: A process perspective. Information & Management 45(1), 31–42 (2008)
19. Heide, J., John, G.: Do Norms Matter in Marketing Relationships? Journal of Marketing 56(2), 32–44 (1992)
20. Henderson, J.C.: Plugging into Strategic Partnerships: The Critical IS Connection Management Review. Sloan 31(3), 7–18 (1990)
21. Kanter, R.M.: Collaborative Advantage: The Art of Alliances. Harvard Business Review 72(4), 96–108 (1994)
22. Kern, T.: The Gestalt of an Information Technology Outsourcing Relationship: An Explanatory Analysis. In: Proceedings of the Eighteenth International Conference on Information Systems, Atlanta, pp. 37–58 (1997)
23. Kern, T., Willcocks, L.: Exploring Information Technology Outsourcing Relationships: Theory and Practice. Journal of Strategic Information Systems 9(4), 321–350 (2000)
24. Kern, T., Willcocks, L.: The relationship advantage. Oxford University Press Inc., New York (2001)
25. Kishore, R., Rao, H., Nam, K., Srinivasan, R., Chaudhury, A.: A Relationship Perspective on IT Outsourcing: A longitudinal study at four companies provides valuable insights about the evolution of IT outsourcing relationships. Communication of the ACM 46(12), 87–92 (2003)
26. Klepper, R.: Outsourcing Relationships. In: Khosrowpour, M. (ed.) Managing Information Technology InvestmentsWith Outsourcing, pp. 218–243. Idea Group Publishing, Harrisburg (1994)
27. Klepper, R., Jones, W.: Outsourcing Information Technology, Systems and Services. Prentice Hall, New Jersey

28. Konsynski, B.R., McFarlan, F.: Information partnerships: Shared data, shared scale. Harvard Business Review 68(5), 114–120 (1990)
29. Lacity, M., Hirschheim, R.: Information System Outsourcing: Myths, Metaphors, and Realities. John Wiley & Sons, Chichester (1993)
30. Lacity, M., Willcocks, L.: An empirical investigation of information technology sourcing practices: Lessons from experience. MIS Quarterly 22(3), 363–408 (1998)
31. Lee, J.N., Kim, Y.G.: Effect of Partnership Quality on IS Outsourcing Success: Conceptual Framework and Empirical Validation. Journal of Management Information Systems 15(4), 29–61 (1999)
32. Lee, J.N., Kim, Y.G.: Understanding Outsourcing Partnership: A Comparison of Three Theoretical Perspectives. IEEE Transactions on Engineering Management 52(1), 43–58 (2005)
33. Levina, N., Ross, J.: From the vendor's perspective: Exploring the value proposition ininformation technology outsourcing. MIS Quarterly 27(3), 331–364 (2003)
34. Loh, L., Venkatraman, N.: Diffusion ofI nformation Technology Outsourcing: Influence Sources and the Kodak Effect. Information Systems Research 3(4), 334–358 (1992)
35. Macneil, I.R.: The New Social Contract: An Inquiry into Modern Contractual Relations. Yale University Press, New Haven (1980)
36. Malone, T., Crowston, K.: The Interdisciplinary Study of Coordination. ACM Computing Surveys 26(1), 87–119 (1994)
37. Marcolin, B., McLellan, K.: Effective IT outsourcing arrangements. In: Proceedings of the 31st Annual Hawaii International Conference on System Sciences, Hawaii (1998)
38. McFarlan, F.W., Nolan, R.L.: How to Manage an IT Outsourcing Alliance. Sloan Management Review 36, 9–23 (1995)
39. Mejias, R., Shepard, M., Vogel, D., Lazaneo, L.: Consensus and Perceived Satisfaction Levels: A Cross-cultural Comparison of GSS and non-GSS Outcomes within and between the United States and Mexico. Journal of Management Information Systems 13(3), 137–161 (1996)
40. Narus, J., Anderson, J.: Consensus Distribution Contributions to Partnerships with Manufacturers. Business Horizons 30, 34–42 (1987)
41. Pearlson, K.E., Saunders, C.S.: Strategic Management of Information Systems. John Wiley & Sons, Inc. (2009)
42. Perrin, B., Pervan, G.: IT Outsourcing Relationship Management and Performance Measurement System Effectiveness. In: 16th Australasian Conference on Information Systems, Paper 54, Sydney (2005)
43. Pfeffer, J., Salancik, G.: The External Control of Organizations: A Resource Dependence Perspective. Harper & Row, New York (1978)
44. Qureshi, M., Kumar, D., Kumar, P.: Modeling the logistics outsourcing relationship variables to enhance shippers' productivity and competitiveness in logistical supply chain. International Journal of Productivity and Performance Management 56(8), 689–714 (2007)
45. Rao, H., Nam, K., Chaudhury, A.: Management of information systems outsourcing: A bidding persepective. J. MIS 12(2), 131–159 (1995)
46. Ring, P., Van de Ven, A.: Developmental Processes of Cooperative Interorganizational Relationships. Academy of Management Review 19(1), 90–118 (1994)
47. Rodriguez, C.M., Wilson, D.T.: Relationship Bonding and Trust as a Foundation for Commitment in US–Mexican Strategic Alliances: A Structural Equation Modeling Approach. Journal of International Marketing 10(4), 53–76 (2002)

48. Sharma, R., Apoorva, S., Madireddy, V., Jain, V.: Best practices for communication between client and vendorin IT outsourcing projects. Journal of Information, Information Technology, and Organizations 3, 61–93 (2008)
49. Simpson, J., Mayo, D.: Relationship Management: A Call for Fewer Influence Attempts. Journal of Business Research 39(3), 209–218 (1997)
50. Skjoett-Larsen, T.: Third party logistics–from an interorganizational point of view. International Journal of Physical Distribution & Logistics Management 30(2), 112–127 (2000)
51. Sun, S., Lin, T., Sun, P.: The Factors Influencing Information Systems Outsourcing Partnership – A Study Integrating Case Study and Survey Research methods. In: 35th Hawaii International Conference on System Sciences, Washington, DC, p. 235 (2002)
52. Swar, B., Moon, J., Choe, Y.C., Oh, J.: Determinants of Relationship Quality for IS/IT Outsourcing Success in Public Sector: A Bilateral Perspective. In: AMCIS 2009 Proceedings, Paper 32. Curran Associates Inc., SanFrancisco (2009)
53. Tian, Y., Lai, F., Daniel, F.: An examination of the nature of trust in logistics outsourcing relationship: Empirical evidence from China. Industrial Management & Data Systems 108(3), 346–367 (2008)
54. Van de Ven, A.H., Delbecg, A.H., Koenig, R.: Determinants of Coordination Modes within Organizations. American Sociological Review 41(2), 322–338 (1976)
55. Vintar, M., Stanimirovic, D.: Evaluation of Impact of Outsourcing on Efficiency of Public Sector Organisations. International Journal of Business and Management Studies 3(2), 211–225 (2011)
56. Waheed, U., Molla, A., Okunoye, A.: Relationship Antecedents in Outsourcing Success: Empirical Investigation of Three Outsourcing Relationships. In: Proceedings of the 2005 Information Resources Management Association International Conference, pp. 626–631 (2005)
57. Willcocks, L.P., Lacity, M.C.: Information systems outsourcing in theory and practice. Journal of Information Technology 10(4), 203–207 (1995)
58. Willcocks, L., Fitzgerald, D.: The outsoucing of information technology: Contracts and the client/vendor relationship. In: The British Academy of Management 1994 Annual Conference Proceedings, Lancaster, pp. 143–156 (2005)
59. Yakhlef, A.: Understanding IT outsourcing. BTJ Tryck AB, Lund (1997)

On the Situated Nature of Designing Knowledge Work Supports Systems

Niclas Eberhagen

Linnaeus University, Department of Informatics, SE-35195, Sweden
niclas.eberhagen@lnu.se

Abstract. There is a need to adopt a situated design perspective in designing computer-based tools that support knowledge work to better understand what it means for users, developers, and stakeholders to approach and capture the tacit knowing within the work context. The design situation is characterized as explorative and iteratively interpreted, a pursuit of the vision of the future system guided by local circumstances. Formal engineering methods, reducing development work to engineering endeavors based on a rationalistic perspective, are not sufficient. The situated design perspective is presented as a conceptual model of the design practice, highlighting its constituent worlds, processes, and relations. The model depicts designing as an explorative and sense-making process, navigating between what is wanted or envisioned and what may be negotiated and discovered. It emphasizes the importance of the artifact being designed as a means to capture, communicate, and discover what is possible in the work context. The model makes clear that the design process is highly situated, that it cannot take place outside the work context because of interdependent relationships. It is designing within the living work context, not design for an objectified one.

Keywords: conceptual model, situated designing, explorative and iterative design, knowledge work context.

1 Introduction

There is a need to provide users and designers, engaged in the development and design of computer-based tools supporting knowledge work, with images and perspectives of such development work that goes beyond those of managerial project plans, milestones, and engineering endeavors. These are needed to better prepare those involved, as well as other stakeholders, with a sound respect of what such work demands and what is to be expected in the planning of such undertakings.

This need stems from two sources, first from the experience drawn from two development projects, presented and discussed in a later section, and then examining the nature of knowledge work. The designing of computer-based tools supporting knowledge work does not only produce artifacts to be embedded in such a practice, but is, in all actuality, also situated within it. Hence, design work comes to share

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 30–44, 2014.

many of the qualities of knowledge work, e.g. the situatedness, tacitness, the socialness, the unstructured nature of such work, and the conditions for how such work is performed. Different actors (users and designers) bring with them their private worlds of engagement into a common world of designing, where transitions of knowledge (design knowledge) occur whenever these worlds overlap. Designing in practice is an ongoing knowing in practice and only understandable within the context of the same.

A conceptual model has been developed which encapsulates the perspective of situated designing in explicit linguistic constructs to highlight its constituent worlds, processes, and relations. Thereby, it provides a means to focus attention and dress debate on what situated designing is, i.e. it provides the very instrument to name and distinguish phenomena. By using explicit concepts such as types of design engagements and processes of knowledge transitions, an operationalization of the model is enabled for studying and exploring other situations of situated designing.

The situated design perspective is presented as a conceptual model of the design practice, highlighting its constituent worlds, processes, and relations. The model depicts designing as an explorative and sense-making process, navigating between what is wanted or envisioned and what may be negotiated and discovered. It emphasizes the importance of the artifact being designed as a means to capture, communicate, and discover what is possible in the work context. It recognizes that final goals are gradually more concretized as the understandings of the work context becomes more articulated and that final specifications are seen as natural stopping points when these converge.

The conceptual model makes visible several important aspects of the design practice. It reveals to those involved, i.e. users and developers as well as other stakeholders, how they are to understand and relate to the design process. It furthers the understanding of what to expect in terms of commitments and efforts that are involved in this kind of design work, including the nature of the roles and responsibilities of those involved. The model makes clear that the design process is highly situated, and that it cannot take place outside the work context because of interdependent relationships. It is designing within the living work context, not design for an objectified one. Thus, it cannot be planned as a pure engineering endeavor, but needs to be viewed as a situated practice.

2 On Knowledge Work

Drawing upon scholarly work on knowledge work ([1, 2, 3]), knowledge in practice ([4, 5, 6, 7, 8, 9]), and on learning in practice ([10, 11, 12, 13, 14, 15, 16]) knowledge work may be characterized as follows.

First, the knowledge work practice is about work that is emergent and non-routine, in some instance highly creative. It may be seen as a response to arising situational conditions and not easily captured in work scripts, and as making do with what is at hand, like “bricolage”, e.g. [17]. This work not only transforms and produces knowledge but also draws upon sources of tacit knowledge for its completion. It is work that is situated, in so far that it is the living context of the work that makes it meaningful, not only to the one participating and engaged in it but also to those contributing to and benefiting from it. It is thus social in that meaning, i.e. how observed phenomena should be interpreted and understood, is socially constructed and negotiated. For the output of the work to be useful and meaningful to others, it must be recognized as such beyond the immediate borders of the practice.

Second, knowledge in the practice is always tacit, a human property whose meaning is socially defined and negotiated, as pointed out above. Such tacit knowledge cannot simply be separated from the practice, captured, packaged and exported as a de-contextualized commodity and applied in another setting effectively. It is only made visible through the ongoing enactment in and the activities within the practice. The dependency upon or inseparability from a living ongoing practice favors a view of knowledge best described as knowing enacted in practice. Such a view summarizes the points made above and at the same time highlights not only the situated nature of knowing practice but also both its tacit and social dimensions. The way to become knowledgeable in and of the practice is to be engaged in its activities.

Third, knowing in practice and learning in practice are two sides of the same coin. They are inseparable or at least tightly coupled to each other, the one conditioning the other. Learning is thus situated in practice and inseparable from the engagement in practice, requiring an active work setting. New understandings and retransformations of older ones cannot be arrived at unless framed within a living context for trying out insights and gaining experience. Work-based learning cannot simply be achieved in a classroom setting, as it is for newcomers a process of being encultured into the practice, observing and gradually trying it out for oneself, and for old-timers a process of honing their skills and being confronted with phenomena at the border their understanding and competence.

3 On Situated Design

The concept of situated design is not an approach to system design as such. Rather, it is a recognizing that the characteristics of the work practice and our evolving understanding of the same dictate, or shape, the conditions for design, and not vice versa. If the work practice is fuzzy and explorative, then so is the process of designing computer-based systems to support it. It is about recognizing that design or development may not be seen as a predefined activity where all the goals are known in advance, but they emerge as the process winds its way forward, defined by the activities performed in the situation.

The concept of situated design is not a widely used term in the literature to describe the above. To help us understand it, one may draw upon scholarly work on situatedness of design [18, 19, 20 drawing upon [21, 22]], on “bricolage” (“designing immediately”) as perspective on situated design [23 drawing upon [24]], and on situated design methodology [25].

As a reaction to the rationalistic view on designing that tries to reduce it into a pure engineering endeavor, Reffat and Gero [18] paint us a different picture of the design process. They view designing as a situated activity, a series of situated acts, where designers interact with their design environment and objects within it, bringing with them prior experience to the particular situation, which cannot be preplanned as designers do not know in advance what will be available at the moment in the situation.

Here, Gero and Kannengiesser [19] have modeled the situatedness in the design situation by introducing three different kinds of environments, or worlds, that recursively interact with one another: the external world, the interpreted world, and the expected world. The external world is the world that is composed of representations outside the agent of design. The interpreted world is the world that is built up inside the agent of design in terms of sensory and perceptual experience, and concepts, i.e. internal and interpreted representation of the part of the external world with which the designer is interacting. The expected world is the world that the imagined actions of the agent of design will produce effects that may be predicted according to current goals and interpretations of the current state of the world. This world is located within the interpreted world, since all goals and expectations may be viewed as interpreted representations of potential future designs.

Also, Gero and Kannengiesser [19] further point out that these three worlds are recursively linked together by three classes of processes: interpretation, focusing, and action. Interpretation means transforming variables sensed in the external world into interpretations of sensory experience, percepts, and concepts that compose the interpreted world. Focusing means taking aspects of the interpreted world and using them as goals in the expected world to suggest actions, which if executed in the external world should produce states that reach the goals. Action entails those resulting effects that bring about a change in the external world, according to the goals in the expected world.

As a reaction to the technology-driven development approach of the rationalistic tradition, Lueg and Pfeifer [25] also seek a better explanation of human cognition based upon the notion of situatedness. Here, human cognition is considered to be emergent from the interaction of the individuals with the environment, i.e. from their involvement in the current situation. Here, Lueg and Pfeifer [25] see the system environment coupling as a prerequisite to cognition that cannot be abstracted away and individuals as situated agents, whose interactions with the environment are dependent upon their cognitive capabilities and experience.

Further, Lueg and Pfeifer [25] describe a methodology for design that take into account both the situation individuals are involved in, without being limited to the task at hand, and the environment in which the task is performed, recognizing the difficulty to predict how the current situation will develop further as new artifacts are introduced.

Situated designing is not a new methodological approach to designing better and more efficient computer-based systems. It is a recognition of the situatedness of everyday work, what that means for the designing of systems supporting such work, and that the design situation and the work practice are intertwined in a continuous cycle of reflection and refinement in approaching the vision of the sought solution.

4 On Development Work

Here follows a brief introduction to the development projects that formed the empirical background to the insights drawn about development work: The Delphi project, and the Dynamite project. For a fuller documentation readers are directed to "Understanding the Designing of Knowledge Work Support Tools as a Situated Practice" ([26]).

The Delphi project began as a joint venture and agreement between Telia, one of Sweden's largest telecommunication companies, and SÄK90/SISY in Sundsvall 1989 under the name of the "Delphi Project". The purpose of the project was to develop a computer-based tool to support security specialists or experts within Telia in their daily work. The support tool as such aimed to serve as a help in organizing information during security analysis work and furthermore serve as a forum for discussion, promoting learning and sharing of meaning through exchanging experience. The support tool also aimed to be able to store large masses of unstructured information and function as a knowledge base for the security specialists for drawing upon experience made with previous similar problem situations. The aim of the development project, in this context, was to provide the users, the security specialists, with a knowledge base regarding the analysis work. Furthermore, it aimed to support them in the managing of the complexity in analysis work by making visible and documenting the analysis work, i.e. serve as a help in organizing information during security analysis work. The knowledge base aimed to provide the users with suggestions for actions and measures by reusing solutions to problem situations, which had been found to be either general in application or proven to have characteristics that had rendered them successful. Further, the knowledge base served as a discussion forum by capturing and representing the analysis work allowing the users to communicate their ideas, insights, and experience with each other. For the possible solutions reached, in resolving a threat scenario, support for testing and simulating is offered in order to see how the costs associated with the different measures evolves.

The MDSS (Maintenance Decision Support System) project represented the work, within work package 12 of the Dynamite (Dynamic Decisions in Maintenance) project, a European Community funded research project. The main objective of the MDSS is to enhance maintenance decision accuracy and cost-effectiveness in order to reduce the economic losses and consequently enhance company profitability and competitiveness. The MDSS supports work of maintenance officers in the production industry by: enhancing accuracy of maintenance decisions though predicting the vibration level (as a condition monitoring parameter) of a component/equipment for the near future, such as at the next planned maintenance action or measuring opportunity, assessing the probability of failure of the component or equipment in question (using machine past data) at need or when its condition monitoring level (e.g. vibration level) is significantly high, and assessing the residual life of a significant component or equipment; simulating and selecting the most cost-effective maintenance solution, i.e. to simulate alternative solutions suggested for a particular problem when all alternatives are technically applicable, and to select the most cost-effective alternative maintenance solution; and identifying and prioritizing problem areas and to assess the losses in the production time through analyzing production process with respect to production time losses and the causes behind them are used to map, follow-up, analyze, and assess the cost-effectiveness of maintenance, and monitoring, mapping, analyzing, following up and assessing maintenance cost-effectiveness, i.e. the contribution or effect of maintenance actions to company profit. The MDSS was developed in close cooperation with researchers of the Linnaeus University and industry representatives from Volvo in Gothenburg, Fiat in Milan, and Goratu in Spain.

The two development projects could be best described as explorative voyages, where the evolving artifact was used as the vehicle for travelling. It was design work characterized as the developers being encultured into the world of the users and their practice to be able to grasp the elusive tacit knowing, to paraphrase Lave and Wenger [14] when they talk about how apprentices learns the tricks of trade from their masters. However, this went the opposite way also as users were brought into the world of developers and technological possibilities and constraints. It could be characterized as design work that was an iterative process of exploration and sense making, much in line of how Berente and Lyytinen [27] see it. It could be characterized as design work that depended on the work situation, where users tried out the proposed design solutions, reflected upon them, as an instrument of learning to better articulate their tacit understanding and wants.

The design work followed no specific methodology, but rather made do with what was available as the drama was played out. Even such situated design methodology as outlined by Lueg and Pfeifer [25], is too optimistic to have served as a “recipe” for success. One cannot make an analysis of the complete working situation unless one has unlimited resources and access of every aspect of that work. The best to hope for is to grasp some aspects, and, just like throwing a pebble into the pond, observe and reflect upon how the waves ripples outwards, i.e. what effects the initial prototype has and learn from that.

5 On Situated Designing

It is far better to view the design situation as not a methodology to follow but rather as process of designing in the situation, where points of delivery may be defined as natural stopping points in the process when visions and design converge in some form of understanding. Towards this end, is proposed a perspective (or conceptual model) of designing that better reflects this, see Fig. 1. The conceptual model illustrates the design situation as a whole, a situated designing world, compromising different interacting or interdependent sub-worlds, in which design activities takes place between all participants.

Here it seems we may be following in the footsteps of Gero and Kannengiesser [19]. However, these worlds are not depicted in the same manner or use the same labels, although there are similarities. The reason for this is that Gero and Kannengiesser modeled these worlds as nested or embedded, one within the other, (cf. [19, p. 380]), and they are not recognized as such in this case. These sub-worlds are all distinct, and entail different perspectives and activities, although they are interdependent. One could think of them as different tunes or parts of the same musical piece, each contributing in their own fashion to the greater harmony, or think of them as different themes of the same discourse.

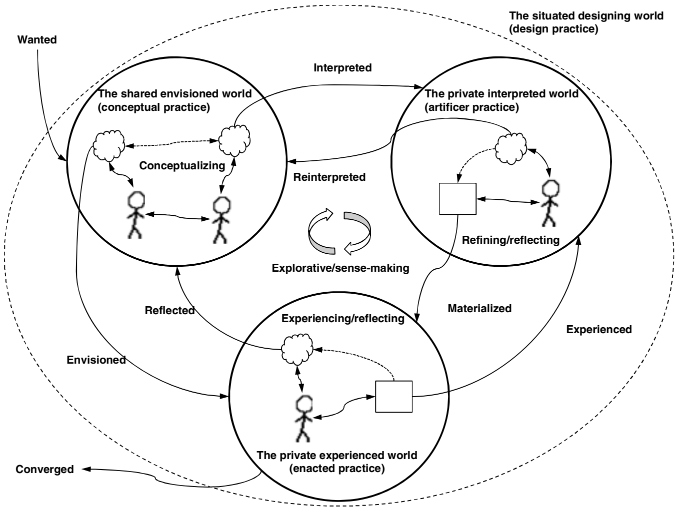


Fig. 1. Worlds of design activities (the design practice) (adopted from [26])

There are many elements within the model, and each will be addressed. However, first there is a need to give a brief explanation of some of them before delving into the sub-worlds and their interdependencies. The clouds represent the idea of what the artifact would or could be, i.e. the vision in our minds. The squares represent the actual artifact. Figure to the far left in each world represent the user(s) and figure to the far right in each world represent the developer(s)/designer(s).

5.1 The Shared Envisioned World

In the shared envisioned world is the "arena" where user and developers come together to make sense and reach a common understanding of where and how they are heading.

It is a world of conceptualization, socialization, and drawing on whiteboards. It is here that the "want" triggers the debate about what may be done to satiate it. As the want enters the world of designing, agents are brought together and this sub-world is nascent as the first theme of the discourse opens up.

At first, it is an arena for enculturation into each other's worlds. Users need to be invited into the world of technological possibilities and constraints, to see what can be done and get an understanding of how it may be realized. Developers need to be invited into the world of the practitioners to build a frame of understanding on which they can lean upon in realizing the design. Both parties being engaged in some form of a legitimate peripheral learning process, in the same sense as Lave and Wenger [14] define it. First observing, as apprentices, then gradually entering further into the domain, as more is mastered, and being able to interact.

Then, gradually as a history of shared explorations into each other's domains emerges as well as a common repertoire of expressive signals to dress thought, it becomes an arena of naming and framing. The want is given flesh to the vision by naming it, giving it an identity, and putting it into different contexts, viewing it with different spectacles, much in the same sense as Schön [21, 22] views designing.

This is also a world of whiteboards. To talk about a design not yet fully realized is an abstract business to be engaged in. The best to hope for is to be able to talk about it through its conceptualizations, i.e. its representations as models and drawings, hence the whiteboards. Sometimes workshops may be held to demonstrate functionality of the prototype system under artificial conditions, at best mimicking some aspect of the practice. Nevertheless, it is still an abstraction not yet come to life. True live experiencing may possible only be achieved in an ongoing working practice.

This is what almost all the projects meeting looked like. It was during these meetings we, the developers, came to have a glimpse of and being gently introduced to the workings of the practice, its concerns and hopes. At the same time, we had the opportunity to suggest technological solutions to meet these concerns and hopes, as well as point out and demonstrate possibilities.

5.2 The Private Interpreted World

This is the private world of developers, or designers, the world of crafting and realization. It is here that the refining and materialization of the artifact takes place, the vision is given physical shape. It is a private world, like the painter working behind his canvas, not easily nor readily accessible or transparent to the users. Designers need to withdraw, with, for the moment, a set of stable assumptions and understandings, in order to mold, shape and try out the evolving artifact.

Into this world is brought the developers interpretation of what the future system will look like, how it will merge with, and function within the work practice. The vision and the evolving artifact are brought into a dialogue, much like the idea of reflective conversation of Schön [22]. The developer creates an imaginary world, built up of what he or she expects to find in the practice, or how he or she perceives to function, i.e. trying to put themselves in the shoes of the users. This imaginary world, a clearly confined micro-cosmos by all means, created with the help of the vision, enables the designers to play, as a user might do, with the artifact, test it, and reflect upon how it works and looks like. Hence, the importance of the enculturation in the shared envisioned world.

Assumptions about the artifact and practice are tested against the vision, serving as the workbench to the developers. This may give rise to tensions due to misconceptions or perceived lacks or gaps of proper understanding. In addition, tension may be brought about in realizing that new possibilities are available enabling to see things differently that the users may or may not have been noticed being caught up in their everyday work life, taking it for granted. All of these could be brought back to the shared envisioned world, as needs for reinterpretation of the vision, to further the fuel of debate there.

5.3 The Private Experienced World

The private experienced world is where the users worked (enacted), experiencing the everyday ongoing business. It is a world that is more or less private to them from the perspectives of us the developers. We may glimpse part of their domain through the feedback we get when they play with our designs, or in the shared envisioned arena where we may become, in part, encultured through building a common vision. We, the developers in the projects, had no other insight into their world. Even if we had workshops and visited their workplace, we could never reach that tacit understanding of what their world was all about, the insights and shared histories, their anxieties and joys, i.e. to walk in their shoes. Unless of course users and developers are the same, all worlds blend perfectly, it is a different matter.

It was to this private experienced world the users brought in the artifact and played with it, confronted it with their visions, and reflected upon it. Immediate reactions to the current conditions could be brought back to us, the developers, as feedback through e-mail and phone conversations. Reflections and explorations having an impact on their vision of the artifact were brought back to the shared envisioned world as new energy to spur further debates and conceptualizations.

5.4 The Input and Output

Let us look at the world of designing as a process with a distinct input and output. The input to the designing world at the outset is best described as something that is "wanted". It is no more specific than that. This "wanted" is the vision, the idea, the intent, the dream of, or longing for something to be accomplished or achieved. Every participant carries an image with him or her, a preconception of some sort. Here, Stolterman [28] has looked at the early vision that developers carry with them into the design process. This "want" is what motivates all the agents, users and developers in this case, to come together and form a world of designing. The energy motivates and spurs productive achievement. In the case of Telia, it was the idea of a support system to help document security analyses and communicate these. That it later grew into something more is a continuation of that "spark" that could be seen as a natural development as more was learned and explored. In this, as the vision manifests itself into a material design it comes to function as means of exploration and a vehicle for learning.

The outcome of the designing process is something that has "converged", the materialized vision. Here the term converge is used much in the same sense as Berente and Lyytinen [27] to point out that natural stopping points in an iterative design process emerge when the wanted and the materialized harmonizes at the moment of observation. It is that which results when a stable artifact encapsulating all that has been learned so far meets the current state of the vision. In development projects, it became that of a fully functioning prototype, a proposal for design by all means, which could be taken as a continuation for exploring the practice again at a later stage. Here the term design proposal is used in the same sense as Bratteteig and Stolterman [29], drawing on Stolterman [30], to denote some form of stable system specification that could be moved into implementation. Of course, other stopping points exist, such as the end of a budget, resources, or time-period. These may coincide with the more natural ones if iterative changes in the design are taken in small steps and discussions in the shared envisioned world are centered on this very fact. The development projects had such end of project budgets. Our development efforts were in tune to what could be achieved up until that specific point in time. As it turned out, our efforts bore fruit. Our proposal for design, the fully functioning prototype, could be picked up again, triggering a new round of debates, if more resources were to be allocated, motivated by further interests in expanding upon and exploring the vision.

5.5 Design World Interactions

As stated earlier, the worlds of designing are interdependent and inter-related. These sub-worlds are not isolated islands of course. We have glimpsed some of this from the previous account of the different worlds. There exist a number of relations, similar to those depicted by Gero and Kannengiesser [19]. It is now time to explore the nature of these inter-relationships.

Between the sub-worlds, there are essentially six relations. Here one must be careful not to view the relations as indicating that some form of objectified entity traverses along these, as something that is transferred across the boundaries between the sub-worlds. Rather, if we were to put the vision of the artifact in the middle of the picture, then the labels point out the state in which the vision of the artifact is currently being in, or how it is perceived.

Let us start with those relations connecting the shared envisioned world and the private interpreted world. From shared envisioned world the developers or designers bring with them their vision of what the artifact is about, the understanding of what the practice is and what the users want. This interpreted vision is what guides and drives the design. As the vision at the same time embodies the designers understanding of what the practice is, it serves as a reference for testing out design alternatives. This is not a one-sided conversation. As the design enfolds and comes to embody the designers understanding it may give rise to tension between the vision and the artifact. Questions may rise due to gaps and needs that were not foreseen. A need for reinterpretation of the vision arises, framed as either questions for clarification or suggestions for further design. These same needs for reinterpretation could rise, as they did on several occasions in the development projects, due to the feedback given to the artifact as users experienced it in their private experienced world.

In between the private experienced world of the users and the private interpreted world of the developers or designers, the vision either belongs in a state of materialization or as experienced. At some point during the designing, the artifact reaches a stage when enough of the interpreted vision of the designers has congealed into physical form and a need for putting it to test arises. Remember, true live experiencing of an artifact may only be achieved once it introduced into the practice. Thus, it is delivered in a materialized form. These points of delivery may be scheduled, as they mostly were in the development projects, or could rise from a need to try out new aspects of the design. Users confronting the artifact, trying play and work with it in situ may come back with immediate feedback as the artifact becomes experienced, leading to fine-tuning, clarifications, and so forth. This interplay between the two sub-worlds may not lead to a questioning of the vision as such, but may give rise to tensions in the private interpreted world suggesting a need for reinterpretation, and may feed the users reflection and questioning of their vision. Often in the case of the development projects this interplay took the form of users asking for technical support, asking how this and that function worked, or reporting errors and bugs, and us developers following up the delivery with questions as to how they received or perceived different aspects of the design.

Finally, in between the shared envisioned world and the private experienced world the vision of the artifact belongs in either a state of being envisioned or reflected. From the shared envisioned world the users will naturally carry with them their idea of what the system will be like, what it may or may not do, their expectations (or anxieties if they have not seen anything yet). As the materialized vision is delivered into their domain and confronted with their everyday on-goings, becoming experienced, it is only natural to expect some form of reaction. This could seed the reflection, learning, and questioning of their vision, and could be brought back to the shared envisioned world at the next meeting there. In addition, the confrontation between their envisioned artifact and the materialized one, could also seed the exploration of their own practice as the artifact could help surface tacit assumptions made explicit through the design. However, what gives rise to these relations is from the point of developers or designers not always transparent. In the development project, we could only catch glimpses of what truly happened in the private experienced world of though interactions in the shared envisioned world and through the interplay between the private interpreted world and the private experienced world.

5.6 The Evolving Artifact

There is one element in the view of the design situation, which needs some further elaboration. It concerns that which is the gravitational point around which all activities centers, that which “traverses” from the wanted and to the converged, the evolving artifact, in between the artist standing in front of the empty canvas with a freshly paint-covered brush and the moment when the painting is uncovered. It is an entity seen to be moving between the sub-worlds, spanning the boundaries of different cultures, domains, and landscapes in much the same sense as Bowker and Star [31] define it. It is an object serving as bridge between technological possibilities and opportunities of the practice. It may seem, as the artifact is really moving between the world, when it is actually the conversation or discourse that changes focus or theme. The artifact, or rather the vision of the artifact, is the focal point for the ongoing dialogue, that which we talk about and with which we travel and explore the universe of the design situation.

Here, Bratteteig and Stolterman [29], drawing on Stolterman [30], have contributed to our understanding of the design situation. They portrayed this "object" as an operative image, which the designers work with. To them, the design situation consists of three levels of abstraction of the future system. Into the design situation designers enter with an initial vision of the future system, and as soon as they start designing the system, it turns into an operative image, in all its possible representational forms, around which all design activities and discussions center. As design stabilizes, and requirements are met, the future system takes the form of a proposal for design (a specification), something that is ready to stand on its own, perhaps as fully functional prototype to be embraced by the users. This is not a linear process, the operative image is both influenced by and influencing the vision. Even the resulting design proposal may influence the vision of the future artifact and trigger further design activities.

Instead of using the concept of Bratteteig and Stolterman [29] of an operating image, one could rather use the concept of Berente and Lyytinen [27] of an iterating artifact. The iterating artifact, the materialization of the evolving vision, is not only the very instrument of capturing, exploring and making-sense of the practice, to meet the needs and wants of the practice with possibilities and limitations of technology, it is also the very discourse, that which we talk about and frames our conversations. It is a vehicle of grasping the multiverse of possibilities as they are played out when our understanding grows, as well as an embodiment of the same. The concept of an iterating (or evolving) artifact reflects better the nature of the design situation. The iterating (or evolving) artifact is in model represented simultaneously as the cloud and the box, the idea and the designed. In the shared envisioned world, it is represented as the idea of the artifact held in common by developers and users.

6 In Conclusion

As has been experienced through the development projects, there is a need for providing a conceptual language to capture better the image of development work that goes beyond those dressed in managerial languages. In managerial languages, the development work is seen as a planned activity, focusing only upon specific points of delivery in a plan to show progress or success. Especially in the second development project, it seemed all too common a practice to regard the project as the plan, with milestones and administrative reports to be produced, all too removed from the situation in which the development work was performed. From the work package group's view, the project was about creativity, problem-solving, trial and error, and so forth. It was much more dynamic, and much more in the here and now. Imagine the frustration when the creative flow was broken up every third month by the needs of the administrative apparatus for control. For sure, there are many good books on system development that try to point out alternative methodologies (e.g. [32]), trying to stress human and organizational issues of learning and knowledge sharing, hiding behind terms such as agile development. Alas, these still emphasize the managerial demands, stressing the project plan primarily. There are no counter images and languages to stem this flow. However, there are alternative images presented. For instance, Bratteteig and Stolterman [29] provide the powerful image of a jazz orchestra when talking about designing in groups, each individual seemingly playing his or her own tune, but all harmonizing as a whole. Nevertheless, such images do not give us the sufficient analytical linguistic tools to uncover and to talk about the mechanisms that make them harmonize in the end.

The laying of the initial model also led to a search for explanations by looking into the context of such development work as experienced in the development projects. Much of what was experienced in the development work of the first project was indeed shared by the characteristics of knowledge work and the nature of knowing in practice. One cannot separate the practice from the designing. Everything is an ongoing knowing (designing) in practice. In designing support systems for a knowledge work practice the designing becomes, so it seems, a knowledge work practice. Designing is here situated in the practice and inseparable from the engagement in practice, requiring an active work setting. New understandings and retransformations of older ones cannot be arrived at unless framed within a living context for trying out insights and gaining experience. It seems that it is only through an active engagement and participation in the practice that we may reach a deeper understanding of what is really going on and why. Here, mere outside observation or retelling would not have lent itself easily to an understanding of what is truly going on, as much of the work depends upon the tacit knowing gained from, or generated as a consequence of, experiencing the effects of one's endeavors.

The role of the system, the artifact (from envisioned to materialized), in designing for a knowledge work practice, is not only as the ends, but also as a means by which much is made sensible and communicated. The "iterating artifact" (the system), the materialization of the evolving vision, is not only the very instrument of capturing, exploring and making-sense of the practice of the users, to meet the needs and wants of the practice with possibilities and limitations of technology. It is also the very discourse, that which is talked about and frames the conversations. Here, it becomes a vehicle of grasping the multiverse of possibilities as they were played out when our understanding grows, as well as a material embodiment of the same. In moving iteratively between users and developers, the prototype system served as a boundary-spanning object, linking the worlds of both users and developers, creating a means for shared understanding.

By adopting the perspective the model provides, we may recognize that characteristics of knowledge work have a bearing upon how developers, together with users, come to approach and capture the rich and tacit knowing of the practice, or part of it, in computer-based tools. The designing of such artifacts in the context of knowledge work should not be viewed as a formal planned effort, as a more traditional engineering approach, which objectifies knowledge of the design situation by assuming that goals may be expressed clearly and understood at the outset of the effort. Rather, it should be viewed as more of an explorative situated process that slowly converges as more is discovered and expressed of the work context, recognizing that knowing in practice means that designing and working blends together, and that these activities are interdependent in a continuous cycle of refining, enacting, and reflecting.

In sum, the conceptual model makes visible several important aspects of the design practice. It reveals to those involved, i.e. users and developers as well as other stakeholders, how they are to understand and relate to the design process. It furthers the understanding of what to expect in terms of commitments and efforts it takes to perform design work, and the nature of the roles and responsibilities of those involved. It makes clear that the design process is highly situated. That it cannot take place outside the work context, but in close interdependent relationship. It is designing within the living work context, not design for an objectified one, and thus, it cannot be planned as a pure engineering endeavor.

References

1. Scarbrough, H.: Knowledge as Work: Conflicts in the Management of Knowledge Workers. Technology Analysis & Strategic Management 11(1), 5–17 (1999)
2. Scarbrough, H.: The Management of Knowledge Workers. In: Currie, W., Galliers, R. (eds.) Rethinking Management Information Systems: An Interdisciplinary Perspective, pp. 474–496. Oxford University Press, Oxford (1999)
3. Schultze, U.: A Confessional Account of an Ethnography About Knowledge Work. MIS Quarterly 24(1), 3–41 (2000)
4. Orlikowski, W.J.: Knowing in Practice: Enacting a Collective Capability in Distributed Organizing. Organization Science 13(3), 249–273 (2002)
5. Nonaka, I.: A Dynamic Theory of Organizational Knowledge Creation. Organization Science 5(1), 14–37 (1994)
6. Nonaka, I., Takeuchi, H.: The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press, New York (1995)
7. Brown, J.S., Duguid, P.: Organizing knowledge. California Management Review 40(3), 90–111 (1998)
8. Wenger, E.: Communities of Practice: Learning, Meaning, and Identity. Cambridge University Press, Cambridge (1998)
9. Newell, S., Robertson, M., Scarbrough, H., Swan, J.: Managing Knowledge Work. Palgrave, Basingstoke (2002)
10. Torraco, R.J.: Integrating Learning with Working: A Reconception of the Role of Workplace Learning. Human Resource Development Quarterly 10(3) (1999)
11. Tyre, M.J., von Hippel, E.: The Situated Nature of Adaptive Learning in Organizations. Organization Science 8(1), 71–83 (1997)
12. Argyris, C., Schön, D.A.: Organizational Learning: A Theory of Action Perspectives. Addison-Wesley, Reading (1978)
13. Lave, J.: Situating Learning in Communities of Practice. In: Resnick, L.B., Levine, J.M., Teasley, S.D. (eds.) Perspectives on Socially Shared Cognition. American Psychological Association, Washington, DC (1991)
14. Lave, J., Wenger, E.: Situated learning: Legitimate peripheral participation. Cambridge University Press, Cambridge (1991)

15. Brown, J.S., Duguid, P.: Organizational learning and communities of practice: towards a unified view of working, learning and innovation. Organization Science 2(1), 40–57 (1991)
16. Orr, J.E.: Talking about Machines: An Ethnography of a Modern Job. ILR Press, Ithaca (1996)
17. Weick, K.E.: Organizational redesign as improvisation. In: Huber, G.P., Glick, W.H. (eds.) Organization Change and Re-design. Oxford University Press, Oxford (1993)
18. Reffat, R.M., Gero, J.S.: Situatedness: A new dimension for learning systems in design. In: Brown, A., Knight, M., Berridge, P. (eds.) Architectural Computing from Turing to 2000, pp. 252–261. University of Liverpool, Liverpool (1999)
19. Gero, J.S., Kannengiesser, U.: The situated function-behaviour-structure framework. Design Studies 25(4), 373–391 (2004)
20. Gero, J.S., Kannengiesser, U.: An ontology of Donald Schön's reflection in designing. International Journal of Design Science and Technology 15(2), 77–90 (2008)
21. Schön, D.A.: The Reflective Practitioner: How Professionals Think in Action. Basic Books, New York (1983)
22. Schön, D.A.: Designing as reflective conversation with the materials of a design situation. Knowledge-Based Systems 5(1), 3–14 (1992)
23. Büscher, M., Gill, S., Mogensen, P., Shapiro, D.: Landscape of Practice: Bricolage as a Method for Situated Design. Computer Supported Cooperative Work 10, 1–28 (2001)
24. Suchman, L.A.: Plans and Situated Actions: The Problem of Human Machine Communication. University of Cambridge Press, Cambridge (1987)
25. Lueg, C., Pfeifer, R.: Cognition, Situatedness, and Situated Design. In: Proceedings of the Second International Conference on Cognitive Technology, pp. 124–135. IEEE, Aizu (1997)
26. Eberhagen, N.: Understanding the Designing of Knowledge Work Support Tools as a Situated Practice (Doctoral dissertation). Linnaeus University Press, Växjö (2011)
27. Berente, N., Lyytinen, K.: The Iterating Artifact as a Fundamental Construct in Information System Design. In: Hevner, A., Chatterjee, S. (eds.) Proceedings of the First International Conference on Design Science Research in Information Systems Technology (DESRIST 2006), Claremont, CA, USA, February 24-25, pp. 237–266 (2006)
28. Stolterman, E.: How System Designers Think About Design and Methods. Scandinavian Journal of Information Systems 3, 137–150 (1991)
29. Bratteteig, T., Stolterman, E.: Design in groups – and all that jazz. In: Kyng, M., Mathiassen, L. (eds.) Computers and Design in Context, pp. 289–316. MIT Press, Cambridge (1997)
30. Stolterman, E.: Designarbetetsdoldarationalitet – En studieavmetodikochpraktikinom systemutveckling (The hidden rationale of design work – A study in the methodology and practice of system development), RRIPCS: Nr 14.91. Department of Informatics, Umea University, Umea (1991)
31. Bowker, G.C., Star, S.L.: Sorting Things Out: Classification and Its Consequences. The MIT Press, London (2000)
32. Avison, D., Fitzgerald, G.: Information systems development, 4th edn. McGraw-Hill, London (2006)

Level of Detail and Understandability of Enterprise Models – Better Understandability through Higher Complexity?

Birger Lantow

University of Rostock, Chair of Business Information Systems,
18051 Rostock, Germany
birger.lantow@uni-rostock.de

Abstract. Enterprise Models are an important tool for IT-Governance. Diverse stakeholders are working with such models. This work presents a contribution to the study of model understandability. The model comprehension of individuals having basic knowledge of business administration and computer science has been measured in an experiment. The level of detail (high/low) has been chosen for the independent variable of the experiment. The study bases on a model that has been derived from the ArchiSurance case study of the OpenGroup. This model is part of the Archi tool[1] published by Bolton University. The results show a higher model understandability at a higher level detail and thus higher model complexity.

Keywords: Enterprise Modeling, EA, Enterprise Architecture, Model Quality.

1 Introduction

Success of enterprises more and more depends on innovation, productivity and process optimization besides a competitive set of products or services. In order to react efficiently and flexibly to the environment, an enterprise must be aware of its structures, processes, and information systems. Enterprise Modeling contributes to the analysis and purposeful development of enterprises. However, there is a different understanding of the term enterprise modeling. According to Fox und Gruninger in [8], an enterprise model is a formal representation of the structures, processes, information, resources, roles, goals and rules. Vernadat [25] sees enterprise models as externalized organizational knowledge. Van der Aalst et al. [24] put business processes into focus, but also include all aforementioned aspects to their process models.

A goal oriented as-is and to-be modeling is done in the course of Enterprise Modeling. Here, various perspectives may be relevant. This depends on the purpose of the modeling project, e.g. process improvement, information system implementation or product portfolio evaluation.

---

[1] http://archi.cetis.ac.uk/

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 45–56, 2014.

Hence, Enterprise Modeling methods define different perspectives. For example, the Enterprise Knowledge Development (EKD) method models enterprise goals, rules, concepts, business processes, actors, and information systems [22]. Therefore, it goes far beyond pure business process modeling. ArchiMate®[2], which is used in this work, does not define a method itself. It is maintained by the OpenGroup and provides a notation that supports a set of perspectives. ArchiMate's structure is aligned with TOGAF®[3]. The ArchiMate core contains Business, Application, and Technology Layer. There are also extensions, e.g. for goal modeling.

Different stakeholders work together in Enterprise Modeling projects. Here, Stirna et al. [22] distinguish two groups. While method experts have knowledge about the method, its process, notations and so on, domain experts have knowledge about the modeled objects. Domain experts may come from various hierarchy levels and functional areas. In order to participate in Enterprise Modeling, all of them need to be able to understand and interpret their specific view of the model and its interface.

This work asks for the influence of the level of detail (cf. Section 2.3) in an Enterprise Model on the model's understandability. It is based on classical paper models. Hence,

Printed representations of the models are used. The focus lies on ArchiMate's Business Layer.

The remainder of this paper is structured as follows. Section 2 discusses related work regarding model understandability. Based on this discussion, section 3 describes the experimental setup of our work. Section 4 analyses and interprets the experimental results. A summary of the main findings and derived possible future research activities can be found in the concluding section 5.

2 Related Work

A lot of work regarding model understandability has been done. For example, the understandability of process models has received great attention [2][7][8][12][13][14][15][18][21][22]. Since process models are a common part of enterprise models, this work has a high relevance for the topic. In the following, different views und understandability and possibilities to measure understandability are discussed in section 2.1. Then previous work regarding influence factors on model understandability is presented in section 2.2. Section 2.3 finally introduces the level of detail as an influence factor and puts it into relation with the presented approaches.

2.1 Understandability and Its Measurement

For a long time, understandability has been seen as a quality attribute of software. It is included in the respective standards (e.g. ISO /IEC 9126). For example, Balzert [3] describes it as the user's effort in order to understand concept and application. Frameworks regarding model quality also see understandability as a relevant attribute. For example SEQUAL [11], a semiotic approach to the description of model quality,

[2] `http://www.opengroup.org/archimate/`
[3] `http://www.togaf.org/`

names the correct interpretation of model intention by stakeholders and by the used tools as aspects of the “Pragmatic Quality”.

Overhage et al. build their 3QM framework [18] on the quality of business process models partly on SEQUAL, using the mentioned „Pragmatic Quality“. 3QM is specified deeper than SEQUAL. Thus, concrete metrics for the assessment of model understandability are defined. However, Overhage et al. limit their work to the consideration of labels within models. This is in order to allow applicability to graphical and textual models as well. The metrics are based on the number of used synonyms and deviations from standard identifiers.

An automated measurement seems to be difficult. Additionally, there are more influence factors of possibly higher relevance according to the authors (cf. section 2.2).

The measurement of the understandability of a concrete model always aims at the model user’s perception. Based on the given definitions, effort (ISO/IEC 9126) or correctness (SEQUAL) respectively has to be determined. Appropriate metrics can be found the framework by Aranda et al. [1]. This framework provides guidelines for the empirical evaluation of model understandability. It names time (effort) and correctness. Here, correctness is measured by the number of correct answers of the user regarding a set of questions about the model. Both measurement approaches are used in this work. Besides them, Aranda et al. also name user confidence concerning the own understanding and perceived difficulty for model interpretation. These metrics represent a subjective assessment of model understandability.

Defining understandability metrics lays the base for an empirical determination and evaluation of influence factors on model understandability. The following section discusses such research approaches and their results.

2.2 Influence Factors on Understandability

In general, not only model characteristics influence the understandability of a model but also the personality of the model user, the environment, and the purpose of the modeling project play a role. While there are studies regarding the influence of personality (e.g. past experiences in modeling in [16]), environment and modeling purpose stay behind at least in the studies regarding the understandability of business process models. However, these factors have to be controlled or captured respectively in empirical studies on model understandability. Furthermore, the influence of the modeling purpose is depicted in the different sub-models and perspectives defined by Enterprise Modeling Frameworks (cf. section 1). Frank puts emphasizes on this aspect in [9]. An overview of further frameworks is given by Chen et al. in [5].

Studies concerning the influence of model characteristics on model understandability lead to normative for model design. The goal is a high model quality regarding this quality attribute. The 7 Process Modeling Guidelines (7PMG) by Mendling et al. [14] are a prominent example. There is empirical proof that compliance to these guidelines has a positive influence on model understandability. In the context of Enterprise Modeling, especially the minimization of the number of model elements and the restriction to at most 50 model elements are relevant out of the 7PMG. The other guidelines have a strong connection to control flow representation in business process models. Additionally, as also shown by Figl and Laue [8], the type of used model elements has an influence on model understandability besides their number. Further studies are related to the representation of models (e.g. Schrepfer et al. in [21]) or on properties of the model graph (e.g. Dumas et al. in [7]). Simple graph properties (number of nodes (of a certain type), node degree, structuredness) are also the base for six of the 7PMG. In general, a higher number of nodes equals a higher cognitive complexity and thus lower understandability. The same applies for metrics like the average node degree of a graph for example (see also [7]). Moody defines in [17] further principles/guidelines for a better understandability of notations. However, these aim at the notation itself and its syntax. Both are not subject to variation in our work.

2.3 Level of Detail and Understandability

In general, model representations that are fitted to the modeling purpose (sub-models/perspectives) cannot completely meet the information demand of model users. The reason lies in the generality of the frameworks that define for example perspectives [5], in the limitations of used modeling tools, and in the effort to generate model views meeting the users demand. Additionally, the information demand may be just unclear.

It can be assumed that there are model elements which do not correspond to the information demand objectively. Having the possibility to remove them from the model would mean a way to improve model understandability. The model would be less complex. An approach to automatically remove model elements is for example filtering by element types as already shown in [6]. Another approach is to cut hierarchies of aggregating relations. This is done by hiding lower levels in the hierarchy and uniting the other relations of lower level nodes on parent nodes. Examples are activities and sub-processes within a business process or sub-divisions within a department. Generally, hiding lower hierarchy levels would mean a higher abstraction. Since the term abstraction is already used in a different context on the field of Enterprise Modeling (e.g. in [5] and [9]), this work uses the term level of detail instead. The level of detail is the higher the more hierarchy levels of aggregating relations are represented within the model

The following hypotheses are resulting based on the discussion:

H1: Model users show a lower error rate when interpreting models having a lower level of detail. (correctness)

H2: Model users need less time when interpreting models having a lower level of detail. (effort)

In the following, the experimental evaluation of these hypotheses is described.

3 Experimental Setting

A questionnaire based experiment has been conducted in order to investigate the influence of the level of detail on model understandability. The candidates had to answer questions regarding their skills on the area of Enterprise Modeling and regarding their comprehension of the given model. Additionally, the time needed to answer the comprehension questions had been measured.

3.1 Used Models

The base model was the ArchiSurance model that ships with the Archi tool[4] provided by Bolton University. The model stems from the ArchiSurance case study. A part (model M1, figure 1) of that model has been selected considering the possibility to create a readable A4-printout. It contains only Business Layer elements. The number of nodes of the corresponding model graph is 36 connected by 54 edges. This results in an average node degree of 3.

In order to vary the level of detail, a second model variant (M2, figure 2) has been created. It hides the lowest process level. The corresponding model graph contains 25 nodes connected by 37 edges. This results in an average node degree of 2.96.

Following the discussion in section 2.2, M2 should be better understandable. If relations would be included that are visualized by edges (see for example [6] fur such visualizations), the expected difference in understandability is even higher. The average node degree of M1 would be 3.72. The average node degree of M2 would be 3.12.

3.2 Questionnaire

The used questionnaire consist of a part for the self-assessment of personal skills on the area of Enterprise Modeling having 5 questions and a part having 12 comprehension questions regarding the given model. Answering was done in the multiple-choice method. All comprehension questions are specific to the selected part of the ArchiSurance model. Correct answers to all questions were possible based on the information contained in M1 and as well by the information contained in M2.

3.3 Participants

The experiment has been conducted independently in two groups. In group 1 (G1) were 11 students of the 4th term and in group 2 (G2) were 8 student of the 2nd term BSc Business Information Systems at Rostock University. Both groups attended basic courses in Business Administration and Information Systems Design. Event-Driven Process Chains (EPC) have been introduced to the students here. Additionally, the concept of Enterprise Modeling has been presented. At the time of the experiment, G1 had attended the courses one year before. G2 just finished the corresponding sessions in their courses. All participants reported that they did not actively work with Enterprise Models previously. Except for one person in G2, all participants were male.

4 http://archi.cetis.ac.uk/

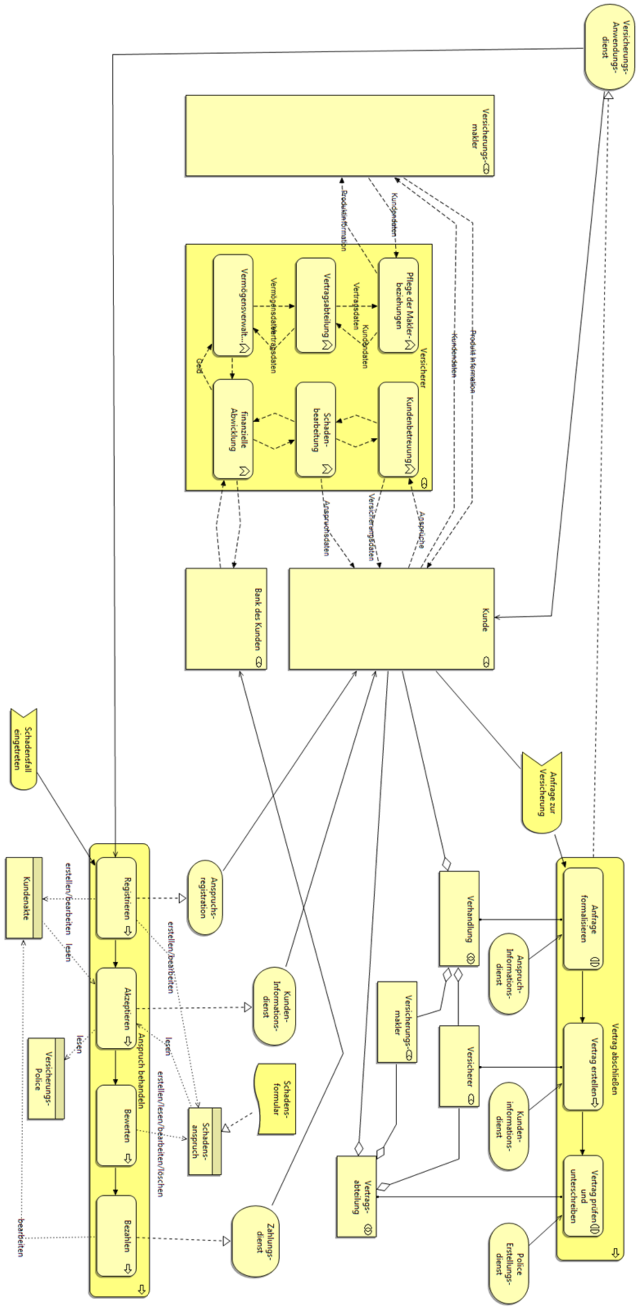


Fig. 1. Model variant M1 (high level of detail)

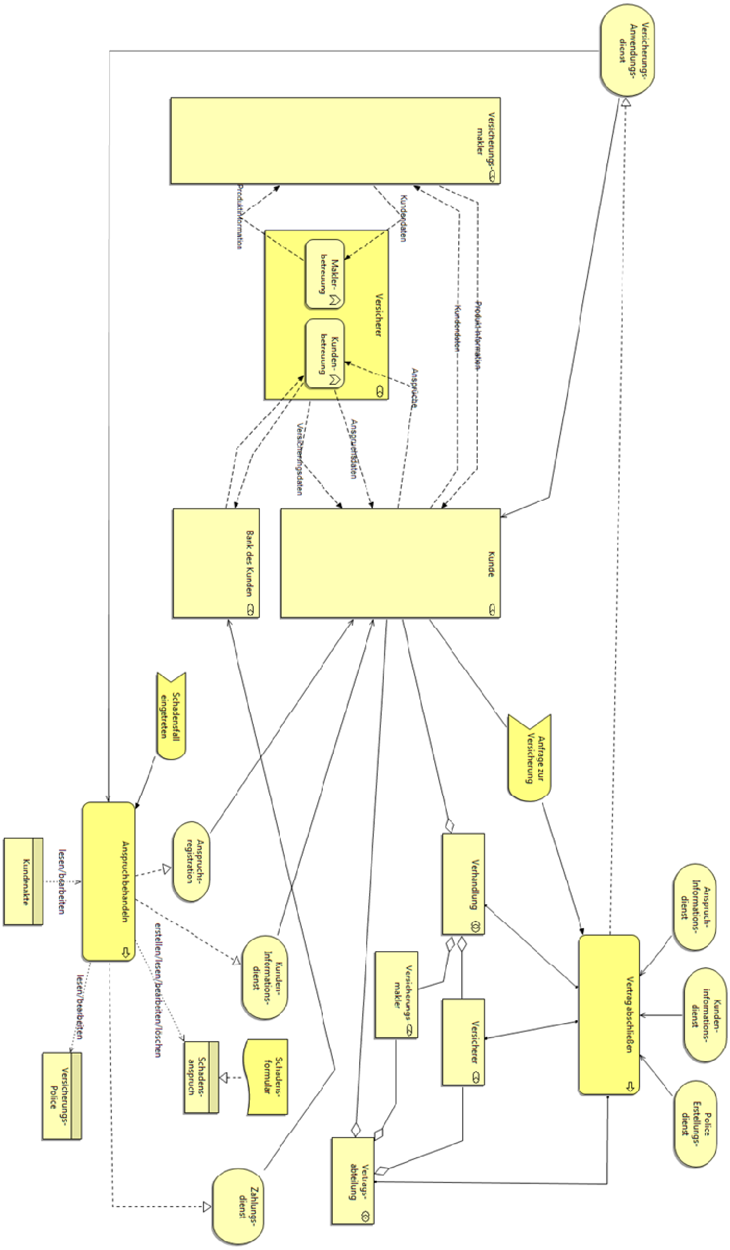


Fig. 2. Model variant M2 (low level of detail)

3.4 Procedure

Each run of the experiment began with an introduction of about 15 minutes into Enterprise Modeling, the ArchiMate notation and the general experimental procedure. Then questionnaires, printed models, and an ArchiMate syntax sheet were distributed to the participants. The model variants – either M1 or M2 - have been assigned arbitrarily. In G1, 6 participants worked with M1 and 5 with M2. In G2, there were 4 participants working on each variant.

In the next step, the participants filled the part of the questionnaire regarding the self-assessment. As soon as everybody completed this, a central timer was started and the participants began answering the comprehension questions. After finishing this part, each participant indicated the elapsed time on his questionnaire.

4 Results

After evaluating the filled questionnaires the following statements regarding the hypotheses can be made.

4.1 H1 Correctness

None of the participants reached the highest possible number of 12 correct answers (see figure 3). There are better results in both groups for the model variant having a higher level of detail (M1).

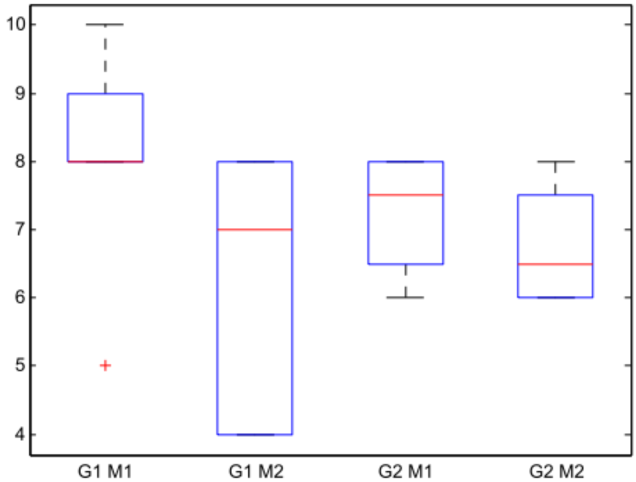


Fig. 3. Boxplot: Number of correctly answered questions

In G1, the median of correctly answered question is 8 out of 12. Except for one outlier with 5 correctly answered questions, all participants have answered at least 8 questions correctly. For M2, the median is at 7 correctly answered questions. There is a wide spread down to the worst participant with 4 correct answers given. The best participant in G1 that worked with M2 had 8 correct answers.

In G2, the median is at 7.5 correct answers for M1 and 6.5 correct answers for M2. It has to be mentioned that replacing the just candidate at the lower end of the spread for M2 with one at the upper end would equalize the difference between the medians for M1 and M2. This is due to the small size of G2.

Overall however, hypothesis H1 is disproved based on the collected data. Users of models with a lower level of detail do not make fewer mistakes in model interpretation. Rather, they make even more mistakes in this special case. Models with a lower level of detail and hence a lower complexity (see section 2.2) show a lower understandability.

4.2 H2 Effort

The user's effort for model understanding has been measured by the time needed in order to answer the comprehension questions. In G1, the median for M1 is at 694.5 seconds and the median for M2 is at 624 seconds. In G2, it is 796.5 seconds for M1 and 636 seconds for M2. The other distribution metrics (quartiles/extreme values) show as well higher values for M1 compared to M2.

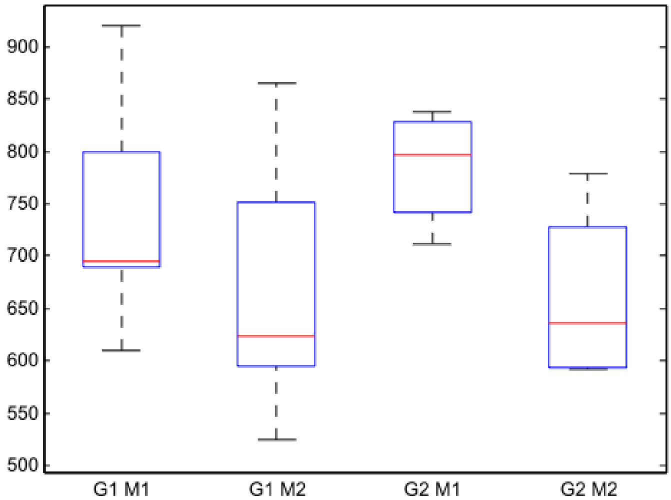


Fig. 4. Boxplot: Time in seconds needed to answer the comprehension questions

Thus, H2 is proved based on the collected data. Users need less time and hence a lower effort for the interpretation of models with a lower level of detail. In consequence, the understandability of models with a lower level of detail would be higher.

4.3 Evaluation and Interpretation of the Results

At first, the experimental results seem contradictory. In terms of the relationship between level of detail and correct model interpretation they are also in contradiction with previous work, e.g. the base theory for 7PMG. In the following, possible reasons for the findings will be discussed.

Besides possible errors in the experimental design which may not have included important influence factors, the small sample size may be an argument against the validity of the found statements. Thus, the found relationship between level of detail and understandability would be due to statistical noise. However, this relationship can be found in both groups independently. This is highly supports it. Additionally, a two-sided t-test including both groups (α=0.1; ν=17; t=1.82) shows, that the null hypothesis of an equal mean number of correct answers for M1 and M2 can be rejected (cf. 4).

Inaccuracies by randomly selected answer option can also be excluded in general. The comprehension questions had at least four answer options each. For a big part of the questions more than one option had to be selected for a correct answer. The probability of choosing a correct answer by chance is considered low. Furthermore, the participants had an “I don’t know” - option that they could select. It can be assumed that correctly answered questions are based on a correct interpretation of the model.

It may also be possible, that the model variant with a lower level of detail was processed less carefully by the participants. More mistakes would be the result. In order to investigate this, the time needed for each model element has been assessed instead of the complete time for reading and interpreting the questions and the model. Since M2 contains less model elements compared to M1 and only the time needed for model interpretation had to be considered, the time for reading, interpreting and filling the questionnaire (questionnaire processing) needed to be deducted. In a post-experiment, this time was estimated retrospectively. In an experiment with 6 participants it was between 100 sec. (1:40 min) and 300 sec. (5:00 min). Taking the extreme values of 300 sec. as a worst case estimate for processing the questionnaire, 4.9 sec interpretation time were needed for each element of M1 and 5.2 sec for each element of M2 as an average for both groups. In consequence, model with a lower level of detail show a lower understandability in terms of effort if an equal number of model elements is assumed. Finally, the experimental results are consistent.

This also shows that the number of model elements has limited suitability for the assessment of model understandability. Furthermore, a modularization of models in order to lower the number of elements per sub-model as proposed by the 7PGM may result in a higher effort for model interpretation if several sub-models are relevant for the model user.

Finally, what is a possible explanation for the experienced effect of a higher level of detail resulting in a better understandability although the model is more complex the same time? Maybe, the additional elements carry context information that supports understandability. The importance of context for model interpretation has already ben emphasized in [6].

5 Conclusion und Outlook

It has been shown, that in certain situations a higher level of detail may improve the model understandability although the models become more complex. Here, global guidelines for model design aim to the wrong direction.

A further validation of the found relationship will be a future task. A goal should be the determination of relevant influence factors or situations that are connected with this relationship. The observation of users interacting with models might be a way to achieve this. This is connected with the demand for a tool support for interactively changes of the level of detail in models.

Acknowledgement. Thanks to Daniel Christen, Florian Karnatz, Felix Timm, and Jörn Wiebring for the vibrant discussion and the support in planning and execution of the experiment.

References

1. Aranda, J., Ernst, N., Horkoff, J., Easterbrook, S.: A framework for empirical evaluation of model comprehensibility. In: Proceedings of the International Workshop on Modeling in Software Engineering, vol. 7. IEEE Computer Society (2007)
2. Ottensooser, A., Fekete, A., Reijers, H.A., Mendling, J., Menictas, C.: Making sense of business process descriptions: An experimental comparison of graphical and textual notations. Journal of Systems and Software 85(3), 596–606 (2012)
3. Balzert, H.: Lehrbuch der Software-Technik. 2. Auflage. Spektrum, Akad. Verl (Lehrbücher der Informatik, 1), Heidelberg (2001)
4. Bleymüller, J., Gehlert, G.: Statistische Formeln, Tabellen und Programme. 8. Auflage. Vahlen (WiSt-Taschenbücher), München (1996)
5. Chen, D., Vallespir, B., Doumeingts, G.: GRAI integrated methodology and its mapping onto generic enterprise reference architecture and methodology. Computers in Industry 33(2-3), 387–394 (1997)
6. Christiner, F., Lantow, B., Sandkuhl, K., Wißotzki, M.: Multi-dimensional Visualization in Enterprise Modeling. In: BIS (Workshops), pp. 139–152 (2012)
7. Dumas, M., La Rosa, M., Mendling, J., Mäesalu, R., Reijers, H.A., Semenenko, N.: Understanding Business Process Models: The Costs and Benefits of Structuredness. In: Ralyté, J., Franch, X., Brinkkemper, S., Wrycza, S. (eds.) CAiSE 2012. LNCS, vol. 7328, pp. 31–46. Springer, Heidelberg (2012)
8. Figl, K., Laue, R.: Cognitive Complexity in Business Process Modeling. In: Mouratidis, H., Rolland, C. (eds.) CAiSE 2011. LNCS, vol. 6741, pp. 452–466. Springer, Heidelberg (2011)
9. Frank, U.: Multi-perspective enterprise modeling (MEMO) conceptual framework and modeling languages. In: Proceedings of the 35th Annual Hawaii International Conference on System Sciences, HICSS 2002, pp. 1258–1267 (2002)
10. Fox, M.S.: Grüninger, Michael: Enterprise Modeling. AI Magazine 19(3), 109–121 (1998)
11. Krogstie, J.: Model-based development and evolution of information systems. A quality approach. Springer, London (2012)

12. Maes, A., Poels, G.: Evaluating Quality of Conceptual Models Based on User Perceptions. In: Embley, D.W., Olivé, A., Ram, S. (eds.) ER 2006. LNCS, vol. 4215, pp. 54–67. Springer, Heidelberg (2006)
13. Melcher, J., Mendling, J., Reijers, H.A., Seese, D.: On Measuring the Understandability of Process Models. In: Rinderle-Ma, S., Sadiq, S., Leymann, F. (eds.) BPM 2009. LNBIP, vol. 43, pp. 465–476. Springer, Heidelberg (2010)
14. Mendling, J., Reijers, H.A., van der Aalst, W.M.P.: Seven process modeling guidelines (7PMG). Inf. Softw. Technol. 52(2), 127–136 (2010)
15. Mendling, J., Reijers, H.A., Cardoso, J.: What Makes Process Models Understandable? In: Alonso, G., Dadam, P., Rosemann, M. (eds.) BPM 2007. LNCS, vol. 4714, pp. 48–63. Springer, Heidelberg (2007)
16. Mendling, J., Strembeck, M.: Influence Factors of Understanding Business Process Models. In: Abramowicz, W., Fensel, D. (eds.) BIS 2008. LNBIP, vol. 7, pp. 142–153. Springer, Heidelberg (2008)
17. Moody, D.L.: The "Physics" of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering. IEEE Transactions on Software Engineering 35(6), 756–779 (2009), doi:10.1109/TSE.2009.67
18. Overhage, S., Birkmeier, D., Schlauderer, S.: Quality Marks, Metrics, and Measurement Procedures for Business Process Models - The 3QM-Framework. Business & Information Systems Engineering 4(5), 229–246 (2012)
19. Parsons, J., Cole, L.: What do the pictures mean? Guidelines for experimental evaluation of representation fidelity in diagrammatical conceptual modeling techniques. Data & Knowledge Engineering 55(3), 327–342 (2005)
20. Poels, G., Burton-Jones, A., Gemino, A., Parsons, J., Ramesh, V.: Experimental Research on Conceptual Modeling: What Should We Be Doing and Why? In: Embley, D.W., Olivé, A., Ram, S. (eds.) ER 2006. LNCS, vol. 4215, pp. 544–547. Springer, Heidelberg (2006)
21. Reijers, H.A., Mendling, J.: A Study Into the Factors That Influence the Understandability of Business Process Models. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 41(3), 449–462 (2011)
22. Schrepfer, M., Wolf, J., Mendling, J., Reijers, H.A.: The Impact of Secondary Notation on Process Model Understanding. In: Persson, A., Stirna, J. (eds.) PoEM 2009. LNBIP, vol. 39, pp. 161–175. Springer, Heidelberg (2009)
23. Stirna, J., Persson, A., Sandkuhl, K.: Participative Enterprise Modeling: Experiences and Recommendations. In: Krogstie, J., Opdahl, A.L., Sindre, G. (eds.) CAiSE/2007. LNCS, vol. 4495, pp. 546–560. Springer, Heidelberg (2007)
24. van der Aalst, W.M.P., ter Hofstede, A.H.M., Weske, M.: Business Process Management: A Survey. In: van der Aalst, W.M.P., ter Hofstede, A.H.M., Weske, M. (eds.) BPM 2003. LNCS, vol. 2678, pp. 1–12. Springer, Heidelberg (2003)
25. Vernadat, F.B.: Enterprise Modeling and Integration: Current Status and Perspectives. Annual Reviews in Control 26, 15–25 (2002)

Performance Modeling for Collaborative Enterprises: Review and Discussion

Barbara Livieri[1] and Mario Bochicchio[2]

[1] Department of Economic Sciences
[2] Department of Engineering
University of Salento, Lecce, Italy
{barbara.livieri,mario.bochicchio}@unisalento.it

Abstract. In the last years, new challenges have arisen for Information Systems (IS) due to the movement towards a society driven by collaboration, whereas Collaborative Enterprises are used as catalysts of competitive advantages. In this context, conceptual modeling can help in the understanding of complex entities and in the development of specific Enterprise IS for intra- and inter-organizational settings. Indeed, the understanding of the domain of analysis is a key-step in the development of IS. In particular, a challenge is posed by the performance management in inter-organizational settings. However, there are still few works that take into account performance related aspects for collaborative enterprises. The aim of this work is to provide a systematic review of literature in order to retrieve, classify, and summarize existing research on performance modeling and to identify areas and opportunities for future research.

Keywords: Enterprise Modeling, Performance, Ontologies, DSML, Information Systems.

1 Introduction

The transition from intra-organizational to inter-organizational relationships characterized the last twenty years of the evolution of enterprises [1, 2]. More recently the increasing importance of Collaborative Enterprises and the growing impact of technologies on businesses drove the transition towards a new phase of trans-organizational relationship, characterized by an increased speed to create value [1, 3]. Indeed, collaborative enterprises, defined as "a number of autonomous organizations working together […] for mutual benefit" [4], are often used as catalysts of competitive advantages.

Nonetheless, research and empirical studies [2, 5] suggest that collaborative enterprises need specialized tools to support performance management and decision-making processes in order to deliver better performance. In this sense, a key role is played by cross-organizational Information Systems, whose development can be supported by enterprise modeling techniques, which are a set of formal and semi-formal able to model, represent and describe important aspects of the structure and of

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 57–71, 2014.

the operational life of an enterprise. According to [6], enterprise modeling can be categorized in five main categories, namely (1) human sense-making and communication, (2) computer-assisted analysis, (3) business process management and quality assurance, (4) model deployment and activation, (5) to give context. For this reason, the research on enterprise modeling has several topics. Some authors focus on the analysis of business processes [7–9], others on the information architecture [10] of firms, some others on the modeling of strategic and organizational aspects as well [11, 12], or of performance indicators, by means of domain-specific modeling languages (DSML) and ontologies.

With the purpose of modeling performance-related concepts, different modeling approaches have been developed relying on domain specific modeling languages as well as on existing and well-known modeling approaches, such ontologies. The aim of these models is to define a shared knowledge on performance measurement, to support the model-driven development of Information Systems Components and, thus, to support firms in the decision-making processes. Although this is a novel field of research, there is an interest towards it from both the business administration perspective and the Information Systems one. Indeed, for several decades the research on business management focused on performance monitoring and control of intra-organizational activities to improve the efficiency of enterprises, whilst, more recently is moving towards the analysis of inter-organizational and trans-organizational relations in order to allow a broader view on enterprise perspectives [1, 13]. Also, from the IS perspective, there is the need for a "*new generation of enterprise systems*" [3] able to face the challenges brought from a collaboration-driven society. However, in order to solve these issues, the IS design should start with "*understanding the business and organizational domain in which the information system is to be embedded*" by means of conceptual models [14].

The aim of this work is to provide a systematic literature review on performance modeling for collaborative enterprises and to classify these articles by means of a comparison framework. This should enable the fostering of future works and the understanding of possible areas of research. Indeed, whilst there is the need for literature on such topic, there are still few works that take these aspects into consideration.

The rest of the work is structured as follows. First, we describe the method used to retrieve, select and classify the articles. In Section III, we analyze the existing literature on performance modeling. Finally, we classifies these works and we draw our conclusions.

2 Method

This study provides a systematic review of literature [15] with a twofold objective:

- to identify, classify, and summarize existing research on performance modeling;
- to identify areas and opportunities for future research.

2.1 Selection Process

For the purpose of this study, we will analyze only works related to the modeling of performance measurement associated aspects and we will exclude works concerning other modeling objects, although we recognize the importance of the other topics of enterprise modeling previously defined. Also, in this work we will not analyze works only related with collaboration issues or on inter-organizational accounting, since reviews on this topics have been already respectively done by [16] and [13].

For the search on databases, as shown in Table 1, we used the following keywords: "Performance measure" or "Indicators, Key Performance Indicators, KPIs" associated with "modeling", "modeling language", framework", "ontology", "ontology model", "semantic model". Databases that were searched include Web of Science, Emerald Insight and Google Scholar. After searching through databases, we performed an iterative improvement based on a) title analysis to exclude non-relevant works; b) bibliography and citations examination on the selected works, to include other relevant articles not reported by databases.

From this process, we retrieved 21 articles that we further analyzed, through the analysis of abstracts and the context of the whole paper, in order to choose the most relevant works. In particular, in the final selection process we excluded works focused more on enterprise architecture than on the modeling by itself and tools oriented to business process monitoring with few modeling aspects of performance [17–21] and on non-formal representation of KPIs [22–26].

2.2 Classification Criteria

All the works analyzed are classified according to four dimensions, which are:

- the modeling method, depending on the use of domain specific modeling languages or ontologies;
- the object of analysis, depending on the focus on Key Performance Indicators (KPIs) or Process Performance Indicators (PPIs);
- the extent of analysis, i.e., whether the model takes into account only indicators or also the firm in a broader sense, with the analysis of goals;
- the level of granularity in the analysis of Key Performance Indicators (KPIs). In particular, we will distinguish between works that focus on single enterprises and works that take into account an inter-organizational setting.

The analysis of the relevant papers identified from the literature in accordance with these four dimensions, enables the categorization of the various models associated with performance measurement. The classification of these works facilitates the association of each model with a modeling technique and with an object of analysis. This also enables models, which have been developed for similar purposes to be compared and be tracked over time and to identify the progress associated with a particular object of analysis.

Table 1. Method adopted for literature review

Keywords	Performance measurement OR Indicators Key Performance Indicators KPIs AND Modeling Modeling Language Framework Ontology Ontology model Semantic model
Database	WoS (Web of Science), Emerald Insight, Google Scholar
N.ro of sources (conferences and journals)	12 articles, 3 of which on journals (all published by springer), 8 on conference proceedings (6 of which published with Springer and 1 with GI-Edition) and 1 working paper.
Results	Nr. of articles retried: 21 Nr of relevant articles: 12

Moreover, in order to classify ontologies we also adopted other criteria related to some ontologies peculiarities, described as follows:

- methodological approach: the criteria, inspired by [16], regards the language used for the development of the ontology and the evaluation process of the ontology. Indeed, it is difficult to distinguish between works designed with an inspiration approach, an induction one, a deduction one or a synthesis one, because the development process of the ontology is rarely made explicit;
- re-use of existing ontologies: one of the main advantages in the use of ontologies is the possibility to re-use existing knowledge in a specific domain [27]. In this sense, it is important to understand if existing ontologies on KPIs re-use other ontologies in order to improve the model;
- reasoning functionalities on KPIs: another peculiarity of ontologies is the reasoning support, which enables to infer new knowledge from the original domain body;
- aim: this criteria regards the purpose for which the ontology has been developed;

– expressivity: each ontology has a different level of expressivity, depending on the Description Logic (DL) family and the specific extensions that are used in the development of the ontology [28].

Finally, we analyze the cross-references among the examined papers in order to assess which works have been stimulated and inspired by which others.

3 Performance Measurement Modeling

3.1 Domain-Specific Modeling Languages

In general, domain-specific modeling languages (DSML) are used to model artifacts in a specific domain of analysis, such as the enterprise one. Some authors focused their attention on DSML for performance measurement. The aim is to offer models able to support the creation and the effective and efficient interpretation of "performance measurement systems […] by providing differentiated semantics of dedicated modeling concepts and corresponding descriptive graphical symbols" [11].

In [29] a framework for business process monitoring through User Requirements Notation (URN), extended with Goal-oriented Requirements Language (GRL) and Use Case Maps (UCM) is proposed. In particular, the authors model the meta-type `Indicator`, that belongs to an `IndicatorGroup` through which it's possible to aggregate KPIs and to offer different views and perspectives to user. Finally, each KPI have a target value (users' expectations), a threshold value and a worst value.

Similarly, Popova and Sharpanskykh [30] developed a framework for modeling KPIs and their relations through dedicated first-order sorted predicate logic-based modeling language, while temporal relations are expressed through Temporal Trace Language (TTL). In order to model indicators, they take into account different views of an enterprise, namely, process-oriented, performance-oriented, organization-oriented and agent-oriented. In the case study, the modeling language is represented by means of circles (for KPIs) and directed edges (for relations). Authors take into account some types of relations among KPIs such as `causing`, `correlated` and `aggregationOf`. As relations among KPIs, tasks, goals, processes, roles and agents, they analyze `is_defined_over`, `is_based_on`, `uses`, `has_owner`, `measures` and `env_influence_on`. The authors focus their study only on processes; however, they do not say how KPIs are calculated and do not propose a specific DSML.

With similar intentions in [31] a Business Intelligence Model (BIM) is used in order to model the strategy, the related goals, the indicators and the potential situations (Strengths, Weaknesses, Threats and Opportunities). In particular, for the modeling of KPIs and their relations the Semantics of Business Vocabulary and Business Rules (SBVR) proposal is used. The SBVR definition is then translated in OCL4OLAP, thus allowing the query of models with a multidimensional structure. In [32], techniques and algorithms to define KPIs metrics expression and value are developed.

In [33] and [34], a method for the design and use of indicator systems is proposed. In more detail, the modeling of the indicator system is achieved by a domain specific modeling language (DSML) named ScoreML. This method should satisfy some requirements such as the possibility to design comprehensive and consistent indicator systems in order to offer users a support on the interpretation of indicators and on the understanding of their rationales. Also, the method should support the representation of indicators at different levels of abstractions, depending on the user type, and enable the construction of software systems. These requirements can be satisfy through a DSML, which can be embedded in enterprise models.

In the model, each `Indicator` has some attributes (e.g., name, description, purpose) and can be `computed from` or be `similar to` other indicators. The users of the system can also define other cause-effect relation between indicators by means of the meta type `CustomizedRelationship`. Also, each indicator can be linked with a `Goal`, a `DecisionScenario`. A `SpecificIndicator`, that is an indicator type, can be associated with a `ReferenceObject`, which can be a `BusinessProcess`, a `Resource` or a `Product`.

In [11] a model for enabling reflective performance measurement, namely MetricM, and a domain-specific modeling language, named MetricML, are offered. After an analysis of the requirements that a method apt to support the reflective design and use of performance indicator should satisfy, the key concepts of MetricML and its semantics are identified.

In this model, the Indicator concepts describes the characteristics of a performance indicator. In particular, the meta-type is described by a formula, a `UnitOfMeasurement`, a `TimeHorizon`, a `sourceOfRawData`, a frequency of measurement, a purpose and an `IndicatorCategory`, which is useful in order to structure large sets of indicators according to user-defined criteria. Furthermore, MetricM takes into account also inter-indicator relations, such as `TransformsRelation` and `IndicatesRelation`, indicator-context relations, through the integration of MetricML with the MEMO modeling languages, and the indicator-goal relation. Indeed, MetricML falls into a comprehensive research work on multi-perspective enterprise modeling (MEMO) [11, 12]. In the framework of a multi-perspective enterprise model, MML (Meta Model Language) has been defined, through which it is possible to model software engineering, social, managerial and economic aspects of the firm.

Even though these works, and in particular [11], offer a broad analysis of performance indicators and of their relations, on DSML and semi-formal frameworks it's not possible to infer new knowledge from the domain.

3.2 Enterprise Ontologies

Nowadays enterprise are entities far more complex than in the past; therefore, it is not easy to manage them. In this frame, there was the need for a "*…a conceptual model* [...that is…] *coherent, comprehensive, consistent and concise…*" [35]. Ontologies can be very effective to represent shared conceptualizations of specific domains [23] and to allow people to reason about sameness and differentness of concepts. Indeed, ontologies can be seen as repositories of concepts, intended as complex information structures tightly interconnected with each other, on which reasoning functionalities can be applied. In more detail, enterprise ontologies are developed and used for several reasons linked with enterprise modeling, such as the development of Management Information Systems and strategic decision support systems, Business Process Reengineering and the construction of Virtual Enterprises. Some of the enterprise ontologies focus their attention on the monitoring of enterprises and, in particular, on the modeling of Key Performance Indicators (KPIs).

In [36], the authors develop an ontological approach for the definition of Process Performance Indicators (PPIs). In the ontology, designed through OWL DL and represented using UML, each PPI has an ID, a target value and a description. PPIs are defined over a `MeasureDefinition` and their value can be calculated by means of:

- Base Measures: the value is calculated without using any other measure. Some base measure are: `TimeMeasure`, `CountMeasure`, `ConditionMeasure`, `DataMeasure`;
- Aggregated Measures: the value is calculated by means of an `aggregationFunction` that enables to get a single value from several values of different instances. Examples of aggregation functions are minimum (`MinAM`), maximum (`MaxAM`), average (`AvgAM`), sum (`SumAM`) or count (`CountAM`);
- Derived Measures: the value is calculated through a mathematical functions, able to
- mbine different `MeasureDefinitions`.

In more detail, the authors define two relationships between measure definitions, i.e., `aggregates`, when PPIs are calculate starting form the same type of measures referred to difference processes, and `isCalulcated`, when PPIs come from different measures. Each measure can affect the other one in a positive way (`is CalculatedPositevely`) or in a negative way (`isCalculated Negatively`); in these cases one measure `dependsOn` another measure or `isDepended`.

However, the authors take into account only process indicators, without, perhaps, considering the relation between goals and KPIs, which is important in order to define a general framework of analysis. A wider range of indicators and the analysis of the related objectives would be indeed useful to assess the overall performance of the firm. Also, there is no formalization of mathematical formulation, so it is not possible to infer, e.g., the relations between indicators or to solve interoperability issues in the case of collaborative enterprises.

In [37], the authors define an ontological model for KPIs in the perspective of Business Process Management (BPM) and, more in detail, of Business Activity Monitoring (BAM), which enables the continuous monitoring of processes' performance. The monitoring is also enabled by Semantic Business Process Management (SBPM), which supports the semantic specification of business objects such as the inputs, outputs, preconditions and postconditions of activities. The same semantic specification can be used to model KPIs in the SBP modeling phase as KPIs are also based on business objects.

The model for KPIs is defined by means of WSML, through which is possible to manage some basic mathematical operations, thus KPIs formula. In particular, each KPI has some attributes, such as name, description, `targetValue` and `analysisPeriod`. The KPI can deviate from the target value in a certain `valueRange`: if the deviation doesn't happen in that range, then an `Alert` is sent. Also, KPIs can have aggregated metrics or instance metrics. On turn, the latter can be:

- `StateMetric`, in which case it evaluates ta condition with a true or false statement;
- `DurationMetric`, in which case it evaluated the time interval between two activities.

Moreover, the instance metrics can be composed by means of operators (max, min, average and sum) in order to compose `aggregate metrics`, which have a `composition expression`.

An interesting work has been done in [38] where an ontology of KPIs with reasoning functionalities for Virtual Enterprises is presented. The ontology, named KPIOnto, is developed within the BIVEE project, which aim is to integrate and share knowledge in innovation projects. In particular, the model enables the definition and manipulation of heterogeneous KPIs calculated in partner firms. The indicators are modeled as instances of the class `Indicator`, which has some data properties (i.e., name, identifies, acronym and definition) and has as subclasses the enterprise area of intervention, e.g., `ProductInd`, `OperationInd`, `CustomerInd` and `CorporateInd`. Also, each indicator can have a `Dimension` and a `Formula`. For the former, a multidimensional model is adopted: a dimension is usually structured into a hierarchy, i.e. a directed tree whose nodes are dimensional attributes and whose arcs represent many-to-one associations between dimensional attribute pairs. Each path in the tree represents a different way of grouping the instances of the class `Indicator`. In more detail, the class `Dimension` can be de-composed in its subclasses such as `OrganizationDimension` and `TimeDimension` and each dimension has different levels of analysis. This structure enables the use of OLAP operation (roll-up and drill-down) on KPIs. Each `Formula` is related to the correspondent `Indicator`, has an aggregation function and is connected to the component values. The mathematical formulation of KPIs and its semantic is represented by means of MathML whilst operators semantics are modeled by means of OpenMath.

As stated from the authors, "*the formal representation and manipulation of the structure of a formula is essential in Virtual Enterprises to check inconsistencies among independent indicator definitions, reconcile indicators values coming from different sources, and provide the necessary flexibility to indicators management*" [38]. To this end, in order to enable reasoning functionalities, the authors use Prolog and refer to XSB as reasoning engine. KPIs formulas, expressed in MathML contents are translated in LP facts. The main reasoning functionalities are the following:

- Formula manipulation: is used to derive relations between indicators, to rewrite formulas or to simplify them;
- Equivalence checking: is used in order to check for duplicates and, whereas found, let the user choose whether to erase them or to allow multiple definition for the same KPI;
- Consistency checking: is used in order to check whether changes in the ontology (new KPIs, updated KPIs, deleted KPIs) affect its consistency;
- Extraction of common indicators: is used in order to extract all atomic indicators needed to calculate a set of indicators.

This work in one of the most complete and relevant for the definition of a KPI ontology for collaborative enterprises. However, authors put much of their focus on innovation processes and not on the firm as a whole. Also, in their model they don't take into account goals or the relations between KPIs, nor their meanings and possible interpretations.

Finally, in [39] the Enterprise Monitoring Ontology (EMO) is defined, with the aim of monitoring value constellations, i.e., systems in which actors collaborate and exchange value objects. In order to develop EMO, the authors extend the Enterprise Ontology proposed by Dietz [40] with concepts related to monitoring aspects. Some of these concepts come from the reification of the corresponding elements of e3value and of the Reference Ontology defined in [41]. EMO is composed by three ontologies, namely the monitoring goal ontology, the monitoring policy ontology and the monitoring metric ontology.

The monitoring goal ontology has three meta-classes: `actor`, `operation` and `object`. Each actor is interested in achieving a `goal`, which can be a `core business goal` or a `monitoring goal`. Goals, which are parameterized through a `metric`, can be accomplished by means of some operations, i.e. `production` and `coordination`, the latter being always connected to production operations. Each operation changes the involved `object`, which can be a `core business object`, which can be defined as a promise to create value or a `monitoring object`, which regards the evidence of the value creation.

In the monitoring policy ontology, the authors restrict the ways in which a goal can be achieved by means of policies. In more detail, a `policy` define what is the role of an actor (the "how") and what permissions does the actor have (the "what"). For each policy an actor can have a specific role, such as `monitor`, `monitoree`, `third-party` or `regulator`, described as subclasses of the meta-class `actor`.

Finally, the monitoring metric ontology qualitatively specifies how goals can be achieved by means of objects. Each `metric` has four dimensions, which are `physical quantity` (e.g., units of measures), `physical quality`, `time` and `space`.

This work is quite extensive, since it's not limited to the analysis of indicators, but takes also into account other aspects, such the role of actors, the policies and the goals, in a value constellation. However, the modeling of the context and of the KPIs is not enough detailed, since, for examples there are several actions that the actors can do in a firm and that don't concerns neither the production nor the collaboration.

Also, in the model there are no reasoning functionalities and it's not clear in which language the ontology has been written.

In general, there are still few works that analyze ontologies of KPIs and a lack of works that simultaneously take into account KPIs, goals and CEs, which are entities far more complex than individual enterprises.

4 Classification Results and Discussion

As depicted in Fig. 1, the works previously analyzed have been categorized according to the modeling method, the object and extent of analysis and the level of granularity. On the horizontal axis, works are subdivided in accordance with the method used in order to model performance indicators: domain specific modeling languages (DSML) or ontologies. On the vertical axis is shown the object of analysis, which can be the performance of processes (PPIs) or the overall performance of the organization (KPIs). Moreover, the shape of the figures represents the goals analysis: the works that take into account the organization's goals are represented with a square shape, while the others are shown as round shapes. Also, the color exemplifies the organizational granularity: models developed for single enterprises are shown as yellow shapes, while models developed for collaborative enterprises are shown as red shapes.

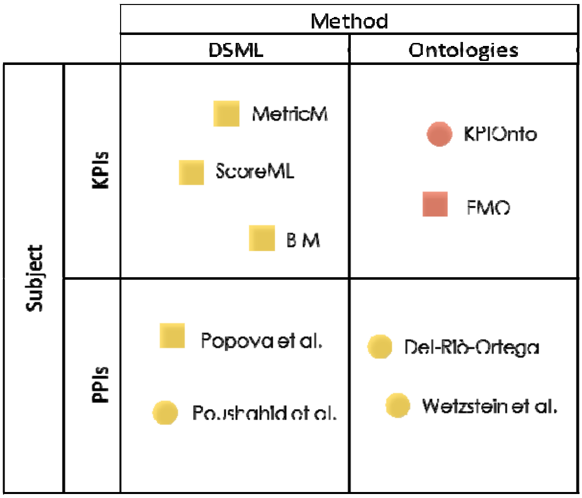


Fig. 1. Classification of the works on performance modeling

From the literature analysis and from the classification, it resulted that there are still few works regarding performance modeling and no works at all that use ontologies in order to conjunctly analyze goals, KPIs and collaborative networks. This could be useful in order to face the challenges coming from inter-organizational performance management. Indeed, as stated in [3] IS should "*enable new forms of* *participation and collaboration, catalyze further the formation of networked enterprises and business ecosystems* […] *ushering in a new generation of enterprise systems*". In this sense, works that consider these issues are particularly useful since in order to develop IS it's important to perform a conceptual modeling phase concerning the domain of analysis [14]. Ontologies can be particularly useful to this aim since complex information systems rely on robust and coherent, formal representations of their subject matter. In this sense, ontologies can provide models of different aspects of a business entity contributing to intra- and inter-enterprise information systems. By committing to the same ontological specification, different applications share a common vocabulary with a formal language and clear semantics. Also, by representing knowledge with a well-established formalism [42], internal consistency and compliance checking can be performed in order to determine content adequacy.

Furthermore, in Table 2 a specific classification for the works that propose ontologies for indicators is offered. From the analysis of the table, it resulted that most of the ontologies developed in this field do not enable reasoning functionalities or, at least, they do not present this kind of result, even though this is one of the main peculiarities of ontologies. Also, the expressivity is not shown, which can be a problem if someone is interested in evaluating "what" the ontology does, even without having access to the whole ontology. In some cases, authors don't even explicit the language in which the ontology is written and in most of the cases the ontology is not available online. Moreover, existing ontologies are rarely re-used, except for the Enterprise Monitoring Ontology, that anyway recalls only ontologies not specifically related to indicators. In this sense, future researches could compare in detail existing ontologies and domain modeling specific languages in order and define a core ontology for indicators, which can then be enriched with domain ontologies. This could be useful also considering that some works could benefit from useful hints coming from similar projects. In Figure 2, the citation between the papers analyzed are depicted.

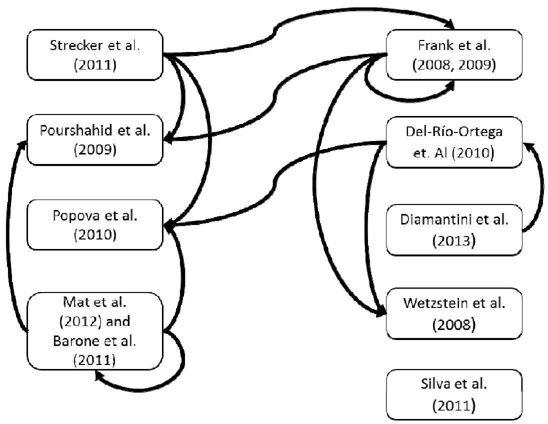


Fig. 2. Citations between selected works

Table 2. Classification of ontologies

Indicators Model	Criteria					
	Methodological approach	Re-use of existing ontologies	Reasoning functionalities	Aim	Expressivity	Object, extent of analysis and level of granularity
Del-Río-Ortega et al. (2010)	OWL DL Evaluation performed through case studies	No	Not declared	Define commonly used PPIs and their relations with business processes	Not declared	PPIs Single enterprise
Wetzstein et al. (2008)	WSML Evaluation through an example	No	Not declared	Integrate the monitoring activities into the semantic business process lifecycle	Not declared	PPIs Single enterprise
KPIOnto (2013)	OWL, MathML and OpenMath, Prolog and XSB as reasoning engine Evaluation through case studies	No	Formula manipulation, equivalence checking, consistency checking and extraction of common indicators	Integrate heterogeneous data in the context of VEs and evaluate common KPIs	Not declared	KPIs Virtual Enterprises
Enterprise Monitoring Ontology (2011)	Language not declared Evaluation through case studies	Enterprise Ontology (Dietz, 2006) Reference Ontology (Andersson, 2006)	Not declared	Provide a framework for the monitoring of value constellations	Not declared	KPIs Goals Value constellations

5 Conclusions

In this work, we present a systematic review of literature with the aim of identifying, classifying, and summarizing existing research on performance modeling and of identifying areas and opportunities for future works. In order to do so, we classified all works according to criteria such as the modeling method, the object and the extent of analysis and the level of granularity. Moreover, we classified ontologies taking into account other criteria such as the methodological approach, the re-use of existing ontologies, the reasoning functionalities, the aim and the expressivity.

From the analysis of the retrieved works according to the criteria previously defined, it resulted that still few works take into account performance modeling with respect to the joint analysis of goals, KPIs and collaborative enterprises. In this field, ontologies can be particularly useful in order to model the domain of analysis, since they enable reasoning functionalities and guarantee a formal representation of the domain. In this sense, it would be useful to have more works that focus on what knowledge can be inferred from the domain, since only few of them show these aspects. Also, a possible direction for future research is the comparison of existing ontologies and domain modeling specific languages in order to define a core ontology for indicators, which can then be enriched with domain ontologies.

Future works will include the broadening of the literature review in terms of number of articles and extent of analysis of them. Also, we will extend the literature analysis in order to account also for works on collaborative enterprises or on single enterprises.

References

1. Bititci, U., Garengo, P., Dörfler, V., Nudurupati, S.: Performance Measurement: Challenges for Tomorrow. Int. J. Manag. Rev. 14, 305–327 (2012)
2. Bititci, U.S., Butler, P., Cahill, W., Kearney, D.: Collaboration: a key competence for competing in the 21st century (2008)
3. FInES Research Roadmap Task Force: Future Internet Enterprise Systems (FInES)-Research Roadmap (2010)
4. Bititci, U.S., Martinez, V., Albores, P., Parung, J.: Creating and managing value in collaborative networks. Int. J. Phys. Distrib. Logist. Manag. 34, 251–268 (2004)
5. Kaplan, R.S., Norton, D.P., Rugelsjoen, B.: Managing Alliances with the Balanced Scorecard. Harv. Bus. Rev., 114–121 (2010)
6. Lillehagen, F., Krogstie, J.: Active Knowledge Modeling of Enterprises. Springer, Heidelberg (2008)
7. Comuzzi, M., Vonk, J., Grefen, P.: Measures and mechanisms for process monitoring in evolving business networks. Data Knowl. Eng. 71, 1–28 (2012)
8. Comuzzi, M., Vanderfeesten, I., Wang, T.: Optimized cross-organizational business process monitoring: Design and enactment. Inf. Sci. (Ny) 244, 107–118 (2013)
9. Pan, Y., Tang, Y., Gulliver, S.R.: Artefact-Oriented Business Process Modeling An Ontological Dependency Approach (2004)
10. Kulkarni, V.: Modeling and Enterprises–the past, the present and the future. Empir. Stud. Softw. Model. 63 (2012)

11. Strecker, S., Frank, U., Heise, D., Kattenstroth, H.: MetricM: a modeling method in support of the reflective design and use of performance measurement systems. Inf. Syst. E-bus. Manag. 10, 241–276 (2011)
12. Frank, U.: Multi-perspective enterprise modeling: foundational concepts, prospects and future research challenges. Softw. Syst. Model. (2012)
13. Caglio, A., Ditillo, A.: A review and discussion of management control in inter-firm relationships: Achievements and future directions. Elsevier Ltd. (2008)
14. Evermann, J., Wand, Y.: Ontology based object-oriented domain modeling: fundamental concepts. Requir. Eng. 10, 146–160 (2005)
15. Tranfield, D., Denyer, D., Smart, P.: Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br. J. Manag. 14, 207–222 (2003)
16. Grubic, T., Fan, I.-S.: Supply chain ontology: Review, analysis and synthesis. Comput. Ind. 61, 776–786 (2010)
17. Davis, R., Brabnder, E.: ARIS Design Platform. Springer (2007)
18. Doumeingts, G., Vallespir, B., Chen, D.: GRAI grid decisional modeling. In: Handbook on Architectures of Information Systems, pp. 313–337. Springer (1998)
19. Chen, D., Vallespir, B., Doumeingts, G.: GRAI integrated methodology and its mapping onto generic enterprise reference architecture and methodology. Comput. Ind. 33, 387–394 (1997)
20. Pedrinaci, C., Lambert, D., Wetzstein, B., van Lessen, T., Cekov, L., Dimitrov, M.: SENTINEL: a semantic business process monitoring tool. In: Proceedings of the First International Workshop on Ontology-supported Business Intelligence, p. 1. ACM (2008)
21. Zeise, N., Link, M., Ortner, E.: Controlling of dynamic enterprises by indicators–A foundational approach. In: Muehlen, M.Z., Su, J. (eds.) BPM 2010 Workshops. LNBIP, vol. 66, pp. 521–530. Springer, Heidelberg (2011)
22. Niedritis, A., Niedrite, L., Kozmina, N.: Performance Measurement Framework with Formal
23. Bertolazzi, P., Krusich, C., Missikoff, M.: An approach to the definition of a core enterprise ontology: CEO. OES-SEO 2001. In: International Workshop on Open Enterprise Solutions: Systems, Experiences, and Organizations, pp. 14–15. Citeseer (2001)
24. Caputo, E., Corallo, A., Damiani, E., Passiante, G.: KPI Modeling in MDA Perspective. In: Meersman, R., Dillon, T., Herrero, P. (eds.) OTM 2010. LNCS, vol. 6428, pp. 384–393. Springer, Heidelberg (2010)
25. Nurmi, A., Moyaux, T., Botta-Genoulaz, V.: Towards semantic performance measurement systems for supply chain management. In: Meersman, R., Dillon, T., Herrero, P. (eds.) OTM 2010. LNCS, vol. 6428, pp. 239–248. Springer, Heidelberg (2010)
26. Matthes, F., Monahov, I., Schneider, A.W., Schulz, C.: Towards a Unified and Configurable Structure for EA Management KPIs. In: Aier, S., Ekstedt, M., Matthes, F., Proper, E., Sanz, J.L. (eds.) PRET 2012 and TEAR 2012. LNBIP, vol. 131, pp. 284–299. Springer, Heidelberg (2012)
27. Uschold, M., Grüninger, M.: Ontologies: principles, methods, and applications. Knowl. Eng. Rev. 11, 93–155 (1996)
28. Baader, F., Horrocks, I., Sattler, U.: Description logics. In: Handbook on Ontologies, pp. 3–28. Springer (2004)
29. Pourshahid, A., Amyot, D., Peyton, L., Ghanavati, S., Chen, P., Weiss, M., Forster, A.J.: Business process management with the user requirements notation. Electron. Commer. Res. 9, 269–316 (2009)

30. Popova, V., Sharpanskykh, A.: Modeling organizational performance indicators. Inf. Syst. 35, 505–527 (2010)
31. Maté, A., Trujillo, J., Mylopoulos, J.: Conceptualizing and Specifying Key Performance Indicators in Business Strategy Models. In: Atzeni, P., Cheung, D., Ram, S. (eds.) ER 2012. LNCS, vol. 7532, pp. 282–291. Springer, Heidelberg (2012)
32. Barone, D., Jiang, L., Amyot, D., Mylopoulos, J.: Composite indicators for business intelligence. In: Jeusfeld, M., Delcambre, L., Ling, T.-W. (eds.) ER 2011. LNCS, vol. 6998, pp. 448–458. Springer, Heidelberg (2011)
33. Frank, U., Heise, D., Kattenstroth, H., Schauer, H.: Designing and Utilising Business Indicator Systems within Enterprise Models-Outline of a Method. In: MobIS, pp. 89–105 (2008)
34. Frank, U., Heise, D., Kattenstroth, H.: Use of a domain specific modeling language for realizing versatile dashboards. In: Proceedings of the 9th OOPSLA Workshop on Domain-Specific Modeling (DSM) (2009)
35. Dietz, J.: Enterprise Ontology: Theory and Methodology. Springer (2006)
36. Del-Río-Ortega, A., Resinas, M., Ruiz-Cortés, A.: Defining Process Performance Indicators: An Ontological Approach. In: Meersman, R., Dillon, T.S., Herrero, P. (eds.) OTM 2010. LNCS, vol. 6426, pp. 555–572. Springer, Heidelberg (2010)
37. Wetzstein, B., Ma, Z., Leymann, F.: Towards measuring key performance indicators of semantic business processes. In: Abramowicz, W., Fensel, D. (eds.) BIS 2008. LNBIP, vol. 7, pp. 227–238. Springer, Heidelberg (2008)
38. Diamantini, C., Potena, D., Storti, E.: A Logic-Based Formalization of KPIs for Virtual Enterprises. In: Franch, X., Soffer, P. (eds.) CAiSE Workshops 2013. LNBIP, vol. 148, pp. 274–285. Springer, Heidelberg (2013)
39. de Alencar Silva, P., Weigand, H.: Enterprise monitoring ontology. In: Jeusfeld, M., Delcambre, L., Ling, T.-W. (eds.) ER 2011. LNCS, vol. 6998, pp. 132–146. Springer, Heidelberg (2011)
40. Dietz, J.L.G.: Enterprise Ontology. Springer, Heidelberg (2006)
41. Andersson, B., et al.: Towards a reference ontology for business models. In: Embley, D.W., Olivé, A., Ram, S. (eds.) ER 2006. LNCS, vol. 4215, pp. 482–496. Springer, Heidelberg (2006)
42. Guarino, N.: Understanding, Building and Using Ontologies: a Commentary to Using Explicit Ontologies in KBS Development. Int. J. Hum. Comput. Stud. 46, 293–310 (1997)

Information Overload: A Systematic Literature Review

Peter Melinat, Tolja Kreuzkam, and Dirk Stamer

University of Rostock, 18059 Rostock, Germany
dirk.stamer@uni-rostock.de

Abstract. Information is essential in our society and for organizations. But the flood of information affects enterprises as well as individuals also in a negative way. This problem is usually referred as Information Overload. This paper summarizes the developments of the past years and gives a prospect to future researches in this field using the method of systematic literature review. A special focus is put on the problems of enterprises with information overload and how these can be solved by using modern approaches.

Keywords: Information Overload, Information Overflow, Systematic Literature Review.

1 Motivation

The central aspect of modern economy is the creation, processing and sharing of information and knowledge [6]. A steady flow of information ensures the quality of products and increases the innovative strength. But the information flow became a raging current as Webster pointed out [30]. The main advantage for an enterprise is no longer just the access to information, but more the access to an adequate information management. A central issue for this is the phenomena of information overload.

Information overload causes several problems, e.g. psychological stress, mistakes in decision making or disregarding of relevant information.

Especially in the past years the importance of this topic increased by the rise of the social networks and mobile access to the internet. These networks allow a steady access to information of the social environment, independent from time or location.

Because of this it is important - from the economical side as well as from the sociological side - to deal with the issue of information overload.

This paper was conducted to get an overview about the recent developments in the field of information overload. For this the following research questions were defined:

RQ1: *How much activity in the field of information overload has there been since 2006?*

RQ 2: *What research topics are being investigated?*

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 72–86, 2014.

RQ 3: *Who is active in this research area?*
RQ 4: *What research approaches are being used?*
RQ 5: *Are there different definitions of information overload? Which?*
RQ 6: *What are the impacts on enterprises due to information overload?*
RQ 7: *What solutions were presented in prospect for these problems?*

The goal is to see what topics are currently researched and to show what aspects of information overload could be promising for future research.

But in the beginning, the idea of *information overload* and its related terms need to be clarified. Then, in the following section, the procedure of systematic literature review is described. The second last section presents the findings of the analysis and put them into relation to the research question conducted above. Finally the conclusion sums up the whole process of research analysis and what future implications can be deduced from it.

2 Theoretical Background: Information Overload

The term overload is commonly understood as the circumstances, where a burden is too strong for the carrying subject. This results in an inability of the subject to work in its full potential or even makes the subject dysfunctional [2, 3]. This interpretation seems to be shared across the most scientific sources and so it should be accepted in this paper, too.

This does not apply to the other important part: *information*. Several interpretations of this term can be made, depending on the point of view and the access to the topic.

A precise definition of information is given by Klaus North, who defines information as data combined with a semantic [19]. In other words: Information is interpretable data. To interpret the information it is important for the receiver to know the semantics. Otherwise it is just some data. This definition fits more to daily experience and common understanding of information.

Another aspect of information is defined by the online business dictionary, where information is defined as a form of data with the following four attributes [1]:

- It is accurate and it is received in time.
- It is organised to fulfil a purpose.
- It has a context, which allows the receiver to assess its meaning and relevance.
- It can help to improve the understanding and decrease the uncertainty.

This approach underlines the economical meaning of information. Especially the focus to fulfil a goal and that it must be received timely indicates a strong will to use the information. But in the common understanding information uses to be information, even if its come to late. The only thing that changes is the value of information. But it seems legit for an economical definition to define the term in this way, because information without a value are useless in their perspective.

If the all these presented definitions for the term *information* show anything, than it is that there are many different aspects of information.

But does information overload refer to different forms of phenomena as well as information? Especially because one of the research questions is about the different definitions of information overload, it can be helpful to determine the phrase before.

Alvin Toffler stated that especially the steady flow of information is not as critical as the speed these information appear for the person [28]. Toffler formulated this in an attempt to illustrate the future of the seventies and especially through the establishment of the internet, it seems to be more true than ever. This is supported by another, more recent statement of Edmunds and Morris. They say, that "*there cannot be many people who have not experienced the feeling of having too much information[...]*" [12]. Hence most people in the scientific community seem to accept the general definition that information overload is the feeling of too much information to be processed for the cognitive capacity of a person [12, 13].

A typical synonym for information overload is the term *information overflow*. Different papers show, that they are used equally and without any distinction [18, 23].

The following section shows how a systematic literature review is structured, how the research was conducted and which papers were selected for a further analysis.

3 Procedure of Systematic Literature Review

The goal of this paper is to review publications, that were released during the past eight years. To do this, the systematic literature review according to Kitchenham [16] was chosen. The advantage of this method is, that it is completely transparent and repeatable, because every step done by the authors, is documented. Another point is, that the research is probably more objective, because the results do not base on one single conference, a specific author or the tendencies of a single search engine. Through the approach the authors have to fulfil the document search in a more comprehensive way. Furthermore this approach seems to fit well for the purpose of this paper, which is to get an overview in the field of information overload.

The review process consists of 6 steps. In the first place the problems and research questions should be formulated, which is already done in the motivation. The next steps is the identification of papers. This means to develop constraints for the research and workarounds for unexpected limitations. Of course, these papers need to be refined some more, so that only relevant papers are part of the review. This paper selection is the third step. For the research about information overload both, the second and third step, are described in the later part of this section. When the final group of papers is defined, data is collected. For the paper at hand this step is not described any further, but in the appendix an extraction of information from the papers can be found. Then the found information is analysed and - according to the questions stated before - answers are given. This leads to the final interpretation of the results. Both steps are topic of the fourth and fifth section, while the fourth is more focused on the analytical part and the fifth is about interpretation. With this short description in mind, step two and three can be started now.

The first point for this is to define the population and describe which parts of the researched paper were examined for the analysis. Because the form of the papers were highly heterogeneous several distinctions were done between those formats. Of course every paper had a title, that was considered in the research, but after the title the papers differ strongly. Some of them follow a research paper style with an abstract and keywords, that were used directly for the research. Others look like article in magazine. Those have no abstract or keywords, but instead a short introduction of the topic, that is already part of the mainmatter. Of course this does not state anything about the quality of papers, but for the literature review they posed a challenge. Discussing this issue, the authors decided to see those introductions in the same way as the abstracts of the regular papers.

To find papers for the literature review, it is necessary to define a search string in order to be able to repeat the literature research getting the same results. The string contains the constraints of the time frame as well as the reference to the topic and synonyms for topic related synonyms. Because the goal of this paper is to review the published articles in the recent years, the time frame was limited to articles since 2006.

As stated before, one synonym for *information overload* is the term *information overflow*. Because of this, both terms were used in the search.

The resulting search string was:

(Title OR Abstract OR Keywords) contain ("information overload" or "information overflow") between 2006-2013

To keep the number of papers manageable, it is necessary to integrate more constrain for the search string. The central restriction was to find sources that fit to the topic and are without charge. The following four sources are included:

– AAMAS (Conference on Autonomous Agents and Multiagent Systems)
– CACM (Communications of the Association of Computer Machinery)
– CHI (Conference on Human-computer Interaction)
– ECIS (The European Conference on Information Systems)

Most of the conferences allow to browse through their content. It shall be stated, that the quality and usability of the search engines could hardly be seen as adequate for this research. One of the problems was, that the search options were not uniform and specific constraints could not be inserted into the engines. For example not all did allow to set a wild card sign (as the "*" in "information over*"). Because of this, it must be considered that several relevant papers could not be found.

To counterbalance the disadvantages, the search process was divided into two steps. In the first step the search string was adapted to requirements of the different search engines.

Furthermore, two different search engines were used in the first step. The first one(`http://cacm.acm.org/search`) brought up 26 papers, while the second one (`http://dl.acm.org/results.cfm`) showed 33 publications. But in the results of the second engine 9 complete magazines were included, which could not be considered as articles. But even after this there were difference in the results. 22 papers were presented by both search engines, while 2 papers were only presented by `dl.acm.org` and 4 only by `cacm.acm.org`. In total 28 papers were seen as relevant after the first research step.

Another problem was caused by the search engine(`http://is2.lse.ac.uk/asp/aspecis/default5.asp`), that was used to find the ECIS articles. This engine did not allow to search just in the title and the abstract and it was not clear if the whole paper or just the title was regarded. Still, 3 research papers were considered as result of the first research step.

For the publications of AAMAS no working search engine could be found. Because of this, it was necessary to search by hand through the 105 available papers.

The engine for CHI (`http://dl.acm.org/results.cfm`) allowed most of the necessary limitations and brought up 101 papers for the first research step.

After the step of population the intervention takes place. Goal of this is to refine the results, that were found during the population. For this research, that meant to put the titles and abstracts(or article introductions) into an Excel Sheet and searched for relevant phrases as *information*, *overload* or *overflow*. When a text contained one or more of these phrases, it was read by the research group and decided, if it is worth to be considered in this analysis. Especially the term *information* is very often represented, but the priority was to find relevant papers. Through this method, it was possible to reduce the number from 247 to 17 final papers. A visualisation can be seen in figure 1.

These 17 papers were considered as suitable to be examined for answering the research questions. For this, a comprehensive analysis was conducted in the following section.

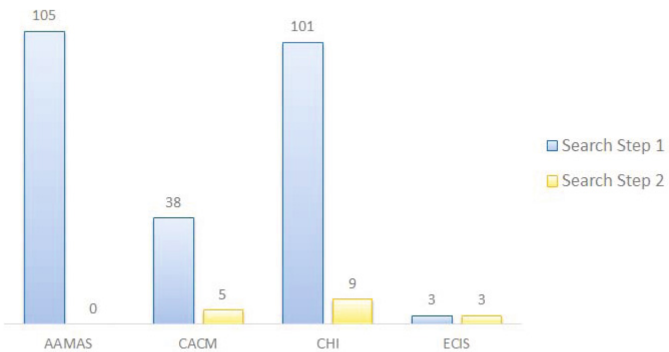


Fig. 1. Number of papers before and after refinement

4 Conduction and Results of the Analysis

In the beginning seven research questions were phrased. In the following part these are connected to the 17 identified papers. For this, the papers were read and analysed to fulfil the tasks.

RQ 1: How much activity in the field of *information overload* has there been since 2006?

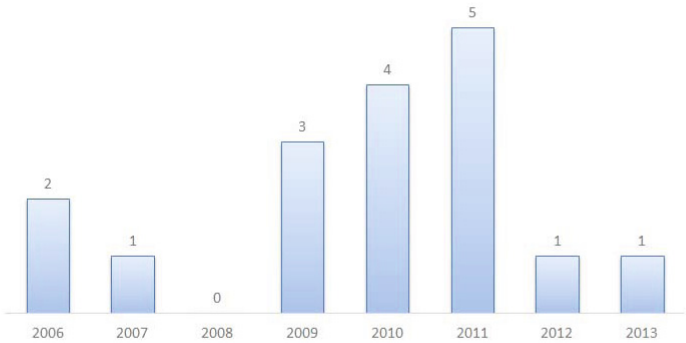


Fig. 2. Distribution of papers in the time from 2006 - 2013

The number of publications in the field of information overload shows that there is little but steady activity (see figure 3.1). In average, 2.125 papers were published per year from 2006 to 2013. The first three examined years showed a straight decrease from two publications in 2006 to zero papers in 2008. After this, from 2009 until 2011, even a little hype could be identified where the number of publications increased. 12 of the 17 researched papers were conducted in this time frame. In the years of 2012 and 2013, this trend could not be continued and the numbers of released papers reduced again to 1 publication per year.

Based on this data, it can be stated that the activity in this field is low and only small number of researchers are active in this field. Nevertheless, the range of topics is wide and different aspects are touched in this field.

RQ 2: What research topics are being investigated?

The field of information overload includes several topics and so do the papers. There are different aspects which should be considered for the classification of the papers.

One form of classification is to determine whether the papers try to identify or communicate a new problem, resulting from information overload, or if they offer an solution for the existing problems. These topics again can be subdivided by the aspect they are focused on, as the organization of people, technological improvements or other methods.

In order to answer this research question, the authors decided for the following five groups:

- **Concepts**, which present usually a general idea how to handle information overload.
- **Solutions**, that aim to implement a general concept or at least construct a framework for an implementation.
- **Evaluations**, which deal with a concept or a solution and try to verify or falsify an approach.
- **Experiments**, that aim, similar as evaluations, to verify or falsify a concept.
- **Examples**, which are illustrating an approach and have always a second group the example applies to.

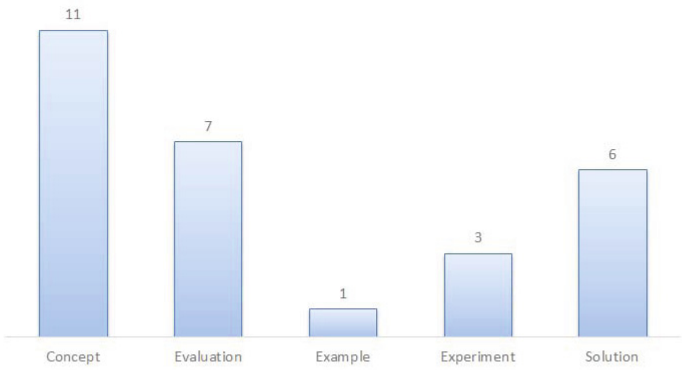


Fig. 3. Topics of the papers

To show the whole coverage of the papers, each paper can be assigned to two of these groups. The distribution can be seen in figure 3.2. Evaluations were usually connected to a solution or a concept they aimed to examine. Only in one case the evaluation was connected to an example [21]. Most papers dealt with a concept (11), or an evaluation (7) of existing concepts or solutions. Six publications were about the creation of a solution, while 3 examined existing solutions or concepts with experiments.

Furthermore, there were several recurring topics. For example, many papers dealt with the questions of filtering information. Most of these suggest to integrate the users in this filtering process, to let them rate the information and filter it based on that [9, 17, 21, 29, 31].

Another topic that was often part of the paper is the processing of information, to make them clearer for recipient [8, 14].

Nevertheless, even with the sophisticated process to search for paper, an essential part of the papers have only a distant relation to the topic of information overload. One for example was about the creation of a pragmatic web [24], another one about the reuse of knowledge in knowledge management systems [9]. Still they were kept for the further steps of analysis, because their input can still valuable.

RQ 3: Who is active in this research area?

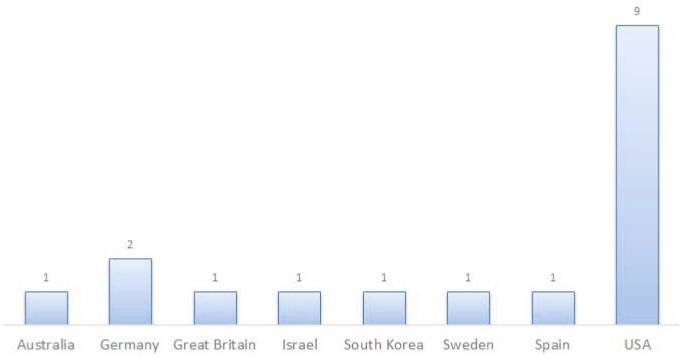


Fig. 4. The distribution of papers according to the nationality of the publishing institution

This questions aims to identify, if there are research networks, with an higher amount of publications in the field of information overload, or if the activity in this field is the same about the whole research community. Another goal is the identification of active research teams by the university and their nationality, to make the reasons for a higher (or lower) activity more transparent. In case of cooperation between different institutes the papers were referred to the institution with the most authors in the paper.

Without considering differences between the first, second and further coauthors, 48 persons were identified as authors. A small number of them work in the industrial sector, but by far the most are employed at universities or other research institutions.

Most of the papers were committed from scientists at the *Palo Alto Research Center*(USA) and the *Naval Posgraduate School* in Monterey (USA), who both have 2 published papers. At the Palo Alto Research Center, the author of both papers was Peter J. Denning, while Ed H. Chi and his team released their papers for the Naval Postgradute School in Monterey.

These numbers show, that there are no scientists who wrote far more papers than anyone else. In general, the distribution is very even among all authors. But this does not apply on the nationality of the publishing institutions. It can be examined that about 50% of the papers were published from institutions in the United States of America, while other nations released only one or two publications. The whole distribution of papers by the nationalities can be seen in figure 3.3.

Several reasons for this can be imagined. One could be that the researched conferences are biased. If, for example, three of the four analysed conference were held in the US, this could be an easy reason for the overweight of US papers. But indeed, the opposite is true. Three of the four conferences are usually held in Europe. Hence the reason for the higher activity in the US is probably determined by something else, for example a special interest in the topic.

Another fact is that selected conferences seems to be focused on Europe and the USA. Only two contributions were sent in from somewhere else (Israel and South Korea). An interesting question for this is if these imbalances had any influence towards the research process.

RQ 4: **What research approaches are being used?**

For a comprehensive analysis it is not only crucial to see the active people and their results, but it is also important to see how they got their results and which methods were used for the knowledge acquisition. During this project, three main groups for research approaches were found:

– Theoretical Work
– Empirical Work
– Case Study

When authors create a new theory and substantiate it by referenced literature, the paper is considered as a theoretical work. The other side of the spectrum is marked by the Case Study. These papers include a practical implementation of ideas and verify or falsify constructed theories. The strong practical alignment usually leads to a less theoretical focus.

But of course, there are also papers which cannot easily be put in any of these two categories because they focus on a theoretical point of view as well as a practical one. This combination can be a theoretical examination that is proved (or disproved) by an experiment. These kind of papers were categorised as Empirical Work.

The distribution of research approaches shows that a majority of researchers uses the form of empirical work for their conduction.

In the section before, it was asked if the national overweight of the USA and the concentration of the western hemisphere affects the form of research. And indeed, there is a shift of papers coming from the US. While about 60% of the theoretical and 62,25% of the empirical work comes from there, they published 25% of the case studies. But considering the small amount of papers, this can be also just a coincidence.

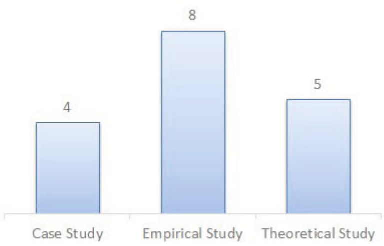


Fig. 5. Distribution of papers according to national affiliation

RQ 5: **Are there different definitions of information overload? Which?**

As stated in the first section, the term *information overload* has a long history and there are several ways to define this term. Five definitions could be examined during analysis process. Some of them were formal and gave a proper description of the phenomena, others were some type of example. One of these formal definitions is given by Whelan and Teigland, when they state:

> ”*The dilemma of having more information than one can assimilate [5] or being burdened with a large supply of information, only some of which is relevant [12], is generally what is meant by information overload.*” [31]

This definition shows different aspects, as the general problem of humans to process more than a limited number of information, and the effort to separate the valuable information from the unnecessary. The same is supported by Otterbacher [21].

A more concise definition is given by Sevinc and D'Ambra, who say

> ”*Information overload occurs as the volume of information received by the individual surpasses their ability to process it*” ([15, 25, 27] according to [26])

Perhaps this definition is not as accurate as the one by Whelan and Teigland, but it shows the main idea of information overload. Another aspect is brought in by Lampe et al., who specify information overload as a problem of the individual [17]. But as stated before, there are also papers which define the term of information overload by example.

As a representative of this second group, Hong defined information overload as the incoming of too many items of communication, ”*e.g. email messages sent by colleagues and friends, news stories related to topics of interest, new tweets posted to Twitter, and status updates in Facebook and LinkedIn*” [14]. In this quasi endless stream of information, people are facing the problem to lose their overview [10, 14].

Another form of overwhelming is the steady change of information, as in the internet or social networks in particular. Cherubini et al. refer to this as *Facebook fatigue* [8].

In general, it has to be stated that all the definitions that were made are quite similar. All of them focus on the overload of impressions caused by the influence of different media. But the social networks brought a new kind of this overload, where the change of information is the critical factor, not only the information flow itself.

RQ 6: **What are the impacts on enterprises due to information overload?**

Several problems caused by information overload can be determined. But which of them especially affect enterprises? Of course those who affect the whole society, have an impact on the enterprises, too.

To start with a very general aspect, information overload causes the victim to need more time for the consumption of the information [7, 10]. These higher time costs of course affect enterprises, too. Furthermore, the people do not concentrate on the given information and block reception. To avoid this, they try to filter the information and try to delegate their responsibilities [10].

On the psychological level, this results in a negative assessment of the own work and competences [11]. And even mental illness can be caused by an overload. Especially the phenomena of burnout is often connected to the steady flow of information [4].

Finally the whole decision making process can be influenced by information overload. People base their decisions on the wrong information or are not able to separate the valuable information from the unnecessary [24]. Additionally there is no time to examine information for its verisimilitude. Hence the made decisions have a lower quality [21, 31].

But not only decision making is affected, even the general work flows suffer from the information overload. The regular check for news disrupts the processes and distracts the people. Because of this, the employees time is wasted and the quality of their work shrinks [26]. A special problem in this field is the *E-Mail Overload*, a form of information overload, caused by the receipt of too many E-Mails [26]. For enterprises, this means that the employees are not able to read every forwarded mail carefully and ignore important details [26, 29].

Furthermore, the overload affects the processes in an enterprise, too. An example for this is the feedback of end-consumer for App developers. The online market places for these applications offer a feedback-portal for the users, where they can suggest software improvement. A problem for the developers, because they have to extract the important facts for future development[20].

RQ 7: **What solutions were presented in prospect for these problems?**

The problems, stated in the question before, are serious and there is no perfect solution for them. Nevertheless, several strategies were conducted and may offer relief from the problems of information overload. These strategies can be divided into strategies for the private environment and strategies for the enterprises.

On the personal level, many problems of information overload are inflicted by the use of social networks (e.g. Twitter or Facebook). Possible solutions for this would be to rate the importance of information by its relevance. One way to capture this relevance could be to use human filters as the personal sphere of contacts in the network [31]. Furthermore, the own postings can be analysed to create a profile of relevant topics each user. This profile again can be used for the relevance assessment [7]. Another form of social filters is the assessment by a jury. This jury can be completely open (as the ranking of customer review in online marketplaces) or exclusive for a specific group of people (as in many online communities). This may help to integrate a social navigation into the portal [21]. Whelan and Teigland state that there are people with a special *information literacy*, who are faster at rating information. Hence it is important to integrate those people as information hubs [31]. This can be an important advantage for enterprises.

Other methods for the organizations can be reduce the number of unnecessary messages. Especially the overload of E-Mails reduces the productivity of employees. An internal policy about E-mail etiquette may helps to reduce this [31]. This can also include to tag E-Mails according to their purpose. Based on this tag, everybody has the opportunity to rate the relevance for themselves [29].

The same procedure applies to other cases as well, for example to show the specific skills of a person to the whole community [17]. Furthermore, this helps to identify experts in a field who act as human filters and enable the reuse of information and knowledge [9]. Another technical solution is offered by Schoop et al., who suggest to make sources of information machine-parsable, to use computers for an adequate supply with information at the right time [24]. Smart push notifications can play an important role in this. These notification publish information, but they act dynamically according to defined *Conditions of Interest*(CoI) and aim to supply the user with valuable information at the right time [10].

But there are not only the technical solutions. The right management plays a key-role to counter information overload. Especially the support of the users self-esteem [26], motivation [4] and commitment management are important methods [11].

After all, it can be stated that the solution strategies depend on four factors [14]:

– The topical content of an information (What?)
– The sender of the information (Who?)
– The way the information is transferred (How?)
– The time of the transfer (When?)

A typical problem of specific solutions is that they do not focus on all the of these facts. E.g., spam guidelines usually only look for the sender of an information. This may work for a specific task, but an holistic application against information overload should include all relevant factors [14].

These answers show, that there is no easy way out of the problems caused by information overload, but there are several approaches that offer promising solutions. Now the last part for this research is about the interpretation of findings for the research questions and give a prospect for future developments.

5 Conclusion

This research started with the definition of the fundamental terms information, overload and information overload. After this, the process of this literature review was explained and it was shown what conferences were examined. During the first steps of analysis 17 papers were considered as relevant for the topic. These papers were used to answer the 7 research questions.

But what are the implications of these answers? And can there be made any prospects or maybe even suggestions for future research projects in the field of *Information Overload*?

For the answer of these question the findings need to be put in a context and connected to each other.

It was found that there is a certain research community in the field of information overload, that publishes in average about two papers per year. Most members of this community work in the United States of America. This bias can be the result of a different cultural perception of the topic or a specific funding by enterprises or governmental institutions.

Furthermore, it has to be stated, that a large majority of papers came from countries that are part of the western hemisphere [22]. This may indicates, that other cultural spheres are not affected as the western culture or, more likely in the opinion of the authors, that there are barriers for researchers from other cultures to participate in the selected conferences. E.g. all conferences are held in Europe or Northern America - this probably reduces the chance for researchers from other parts of the world to participate.

The researched topics were mostly concepts, solutions or validations of those through evaluations or experiments. Papers, that presented technical solutions, were mostly published from institutions in the USA, which may indicates, that the institutes have an higher interest to solve problems of information overload, while others are more interested to phrase theories and evaluate models or solutions.

But in general it can be stated, that the interpretation of the term information overload is widely the same. Each of the found definitions sees information overload as a state, where the input of information is to high for the capabilities of the affected person. The moment this point is reached, depends on the abilities of the person to interact with information.

The problems resulting from information overload are numerous, but especially the higher effort of employees to separate valuable information from others, is a widely recognized issue. This stress even leads to mental health issues.

Solutions for these problems can be technologies or organizational structures, which help to filter incoming information or help people to identify interesting information. There are many forms of these solutions. But most papers only concentrate on one of these aspects. Hence they were focused either on the right management of people affected by information overload or the technology these people use. But no paper combined management and technology aspects to a comprehensive method or showed how these approaches can be combined.

There are no disputes about the meaning of information overload and the problems are defined and accept throughout the whole community. Still there is just a little group of active researchers. This group need to grow, especially because the problems caused by information overload are not going to be easier. As already stated, it seems to be necessary to combine technological approaches with management approaches to formulate a more comprehensive model to counter information overload.

Acknowledgments. The work presented in this paper was supported within the project KOSMOS (Konstruktion und Organisation eines Studiums in Offenen Systemen) funded by the BMBF (Federal Ministry of Education and Research, Germany) and the European Social Funds of the European Union.

References

[1] Business Dictonary.com (December 2014), `http://www.businessdictionary.com/`
[2] Cambridge Dictornaries (January 2014), `http://dictionary.cambridge.org/`
[3] Oxford Dictornaries (January 2014), `http://www.oxforddictionaries.com`
[4] Beverung, D., Wittchen, M., Becker, J.: Where are the participants? including motivational aspects into theorizing and design in is research. In: ECIS (2011)
[5] Butcher, H.: Meeting Managers Needs. Aslib (1998)
[6] Castells, M.: The rise of the network society: The information age: Economy, society, and culture, vol. 1. Wiley. com (2011)
[7] Chen, J., Nairn, R., Chi, E.H.: Speak little and well: Recommending conversations in online social streams. In: CHI (2011)
[8] Cherubini, M., Gutierrez, A., de Oliveira, R., Oliver, N.: Social tagging revamped: Supporting the users need of self-promotion through persuasive techniques. In: CHI (2010)
[9] Delen, D., Al-Hawamdeh, S.: A holistic framework for knowledge discovery and management. CACM (March 2009)
[10] Denning, P.J.: Infoglut. CACM (2006)
[11] Denning, P.J.: The profession of managing time. CACM (2011)
[12] Edmunds, A., Morris, A.: The problem of information overload in business organisations: A review of the literature. Int. J. Inf. Manag. 20(1), 17–28 (2000)
[13] Eppler, M.J., Mengis, J.: The concept of information overload: A review of literature from organization science, accounting, marketing, mis, and related disciplines. The Information Society 20(5), 325–344 (2004)
[14] Hong, L., Convertino, G., Suh, B., Chi, E.H., Kairam, S.: Feedwinnower: Layering structures over collections of information streams. In: CHI (2010)
[15] Jones, Q., Ravid, G., Rafaeli, S.: Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. Information Systems Research (2004)
[16] Kitchenham, B.: Procedures for performing systematic reviews. Technical report (2004)

[17] Lampe, C., Johnston, E., Resnick, P.: Follow the reader: Filtering comments on slashdot. In: CHI (2007)
[18] Lee, D.H., Brusilovsky, P.: Fighting information overflow with personalized comprehensive information access: A proactive job recommender. In: Third International Conference on Autonomic and Autonomous Systems, ICAS 2007, p. 21. IEEE (2007)
[19] North, K.: Wissensorientiere Unternehmensführung: Wertschöpfung durch Wissen 5. Auflage. Gabler Verlag (2011)
[20] Oh, J., Lee, J.-G., Kim, D., Song, J., Lee, U.: Facilitating developer-user interactions with mobile app review digests. In: CHI (2013)
[21] Otterbacher, J.: "helpfulness" in online communities: A measure of message quality. In: CHI (2009)
[22] Porter, R.E., McDaniel, E.R.: Communication between cultures, 8th International edn. Wadsworth Publishing Company (2013)
[23] Savolainen, R.: Filtering and withdrawing: strategies for coping with information overload in everyday contexts. Journal of Information Science 33(5), 611–621 (2007)
[24] Schoop, M., de Moor, A., Dietz, J.L.G.: The pragmatic web: A manifesto. CACM (2006)
[25] Schultze, U., Vandenbosch, B.: Information overload in a groupware environment: Now you see it, now you dont. Journal of Organisational Computing and Electronic Commerce (1998)
[26] Sevinc, G., D'Ambra, J.: The influence of self-esteem and locus control on perceived e-mail overload. In: ECIS (2010)
[27] Speier, C., Valacich, J.S., Vessey, I.: The influence of task interruption on individual decision making: An information overload perspective. Decision Sciences (1999)
[28] Toffler, A.: Future shock. Amereon Ltd., New York (1970)
[29] Tungare, M., Perez-Quinones, M.A.: You scratch my back and i'll scratch yours: Combating email overload collaboratively. In: CHI (2009)
[30] Webster, F.: Theories of the information society. Routledge (2007)
[31] Whelan, E., Teigland, R.: Managing information overload; examining the role of the human filter. In: ECIS (2011)

Developing IT-Enabled Dynamic Capabilities: A Service Science Approach

Patrick Mikalef

Ionian University, Department of Informatics, Tsirigoti Square 7,
49100 Corfu, Greece
mikalef@ionio.gr

Abstract. The transition from a relatively stable goods-based economy to an intangible and highly turbulent service-based economy has necessitated the development of an appropriate theoretical perspective to explain value creation and sustained competitive advantage. Service science has been put forth in order to address this changing view and requires firms to fundamentally reconsider the means by which value is derived. Firms that foster the service logic are increasingly dependent upon Information Technology (IT) to enact their operations and deliver value propositions. Despite significant investments in service-oriented technologies, the prevailing research view regarding management of IT has not been in adherence with the principles of service science or the dynamism of the environment. The view on IT still follows a resource-based logic, in which a competitive advantage is seen as being a result of owning a unique bundle of resources. Literature to date has placed little attention on the IT-enabled capabilities that firms must develop in order to remain competitive in conditions of high environmental turbulence. Building upon this need of conceptually reframing the management of service-oriented technologies which is reinforced by recent calls of the IT management literature, we use service science principles to propose an alternative perspective of IT-enabled capabilities. We ground our developments on the Dynamic Capabilities (DCs) theory since it provides a theoretical basis for explaining resource renewal and competitive survival in highly turbulent environments. Hence, we propose four IT-enabled dynamic capabilities: IT-enabled sensing, IT-enabled coordination/reconfiguration, IT-enabled learning, and IT-enabled integration. We argue that these IT capabilities constitute the prime source of sustained competitive advantage for service-oriented enterprises operating in conditions of moderate to high environmental uncertainty.

Keywords: Service Science, Dynamic Capabilities, Services Computing, IT-Enabled Dynamic Capabilities, Environmental Turbulence, Service-Oriented Architectures.

1 Introduction

The global economy is currently shifting from a goods-oriented economy to an economy in which value creation is primarily dependent upon services [1]. This is

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 87–100, 2014.

reflected in reports of the Organization for Economic Cooperation and Development (OECD), according to which, services currently make up approximately 70% of the aggregate production and employment world-wide [2]. Therefore, it is only natural that companies place a great deal of importance on ways to provide innovative and high quality services to their customers or business partners. This radical shift to a service-dominant economy not only affects companies of the service sector, but also impacts companies in the manufacturing sector, who progressively rely on service operations in order to remain competitive and increase profitability.

A service is defined as the application of competences (knowledge and skills) by one entity for the benefit of another [3]. The value that is derived from services is a result of the interactions between entities that are referred to as service systems. Service systems are configurations of systems that allow for dynamic value co-creation [4]. These systems usually involve people, technology, organizations, and shared information. The notion of thinking in terms of services has been developed in the so called *service science* which aims to *"catalog and understand service systems, and to apply that understanding to advancing our ability to design, improve, and scale service systems for practical business and societal purposes"* [4].

The transition to a service-dominant economy has lead academics and practitioners to reconsider traditional means of Information Technology (IT) in order to facilitate and enable service provisioning in a more agile and scalable manner. The service-oriented paradigm is gaining ground in an increasing amount of companies. It leverages technology in the form of IT Services, in response to the changing business environment, allowing for greater agility, flexibility, and integration with customers and partners. These IT services are enacted through web services, and service-oriented architectures, which are not limited to technological standards, but also reflect a new way of thinking. Firms that have adopted this servitization logic to their business processes as well as their IT infrastructure are referred to as service-oriented enterprises [4].

Despite substantial investments in IT infrastructure in order to support the service-oriented enterprise, a large number of companies are still struggling to leverage them in order to create value and gain a competitive edge. This can be attributed to the fact that it is still not clear how service science can address the managerial aspects of service-oriented technologies [4]. Although there has been much literature on the required organizational capabilities that firms must foster in order to induce change in dynamic environments, the impact of IT capabilities is largely examined under a static prism. This means that IT is examined either as a resource that firms possess (infrastructure, human skills, and relational capital), a competence of the business and IT departments (e.g. IT assimilation, IT planning), or as an operational IT-enabled capability (e.g. IT-enabled automation, support for logistics) [5]. What is lacking in empirical research is a comprehensive view of how through IT-enabled capabilities firms can deal with change in the competitive landscape. In order to understand how value is created in service oriented enterprises it is imperative to examine the enabling capabilities of IT as well as their interdependencies.

Through this perspective it is possible to explore the predominant questions posed in service science and information systems strategy literature about how value is created in service enterprises under conditions of increased environmental uncertainty. To do so, we adopt a dynamic capabilities approach, and attempt to define a set of IT capabilities that firms must cultivate in order to remain competitive. In contrast with past studies, we concentrate on the role of IT-enabled dynamic capabilities, and advocate their necessity in service-oriented enterprises. The need of developing different types of dynamic capabilities has been repeatedly stressed in literature [5], [6]. Furthermore, the Management Information Systems (MIS) research calls for the development of methods that examine the co-evolution of physical and social technologies [7]. In service-oriented enterprises, the need to concurrently evolve business and IT is driven by their fused nature. Since IT is used as a vehicle for enacting business operations, the capabilities needed to adapt a firms` competitive strategy requires the effective management of IT. In addition to developing and describing IT-enabled dynamic capabilities for the service-oriented enterprise, we elaborate on how these capabilities are associated and under what circumstances they may lead to performance gains.

The rest of the paper is structured as follows. In section II, we analyze the concept of service-dominant logic and introduce the means by which services are enabled through IT. Section III describes the foundations of the Dynamic Capabilities (DC) theory, and how it complements the Resource Based View (RBV) of the firm. Based on the analysis performed in the two prior sections, section IV introduces the proposed IT-enabled dynamic capabilities and discusses how literature, particularly in the field of MIS, stresses their importance. We then proceed to explain under what conditions these DCs will lead to performance gains, and pinpoint their inter-dependencies. In closing, we discuss the implications for academic literature and practitioners that this view raises, and how it alters the current view of IT service management.

2 The Service Economy

The transition to a service economy has lead academics and practitioners to reconsider the paradigm by which such value systems should be examined. Service enterprises rely on the idea that the timely application of competences in order to enhance value propositions is the fundamental basis for a sustained competitive advantage. In order to capture the spectrum of service enterprises and how they create value and remain competitive in the turbulent business environments, the development of a new discipline was deemed as a necessity. In the past decade, the concept of service has been redefined from merely an intangible good, to encompass multidimensional theoretical underpinnings, such as *"service-dominant logic"* (S-D logic), *"service science, management, engineering and design"* (SSMED), or simply *"service science"* [8]. In the following subsections we briefly present the underlying principles of these perspectives.

2.1 Service Science

Service science is an interdisciplinary field that combines organization and human elements, with business and technological understanding, to categorize and explain the many types of service systems that exist, as well as how service systems interact and evolve to co-create value [9]. The main idea of service science is to explain how these service systems, which are defined as *"configurations of people, technology, and value propositions connecting internal and external service systems and shared information"*, interact and lead to performance realizations.

Defining and mapping out the process by which new services are developed within service systems is imperative in fast paced domains. Although service science does not set down strict guidelines on constructs or dimensions with respect to empirical studies, it does impose a particular way of thinking. Service science is built on the idea that a service is considered as the fundamental basis of exchange. Hence, it is what is of value for a firm and derives from the application of competencies on resources. The second consideration is that resources (operand resources) that can be easily replicated do not constitute a basis for competitive advantage. Rather, capabilities (operant resources), such human experience, un-codified knowledge, and synergistic networked embeddedness, - or else, are the basis for a competitive advantage. It is through the applications of these capabilities that service systems can offer new value propositions through revamped services. Operant resources can take many forms based on their purpose and area of application [10]. The third premise is that companies cannot deliver value through services *per se,* but rather, value is co-created through the interactions with external systems such as clients or other collaborators. Thus, it is of great significance for all participants involved in the realization of a service to contribute, henceforward the term value co-creation.

2.2 Service Computing

The service-dominated economy and the corresponding service science paradigm require enterprises to be more agile, while at the same time to be capable to forge close ties with customers and partners. This need for flexibility while simultaneously being flexible in operations is reflected both at a strategic and at a technological level [11]. In response to these new challenges, firms need to reconsider the role of IT, and how it can be leveraged to overcome them. Poorly architected IT infrastructures may drastically hinder the ability to react to dynamic changes [12]. Additionally, there is a growing consideration that firms should operate as value networks rather than value chains, a notion which central within the service science discipline.

In congruence with the redefined requirements and views on the modern enterprise, IT should be fundamentally reassessed. To this end, service-oriented technologies such as web services, Service Oriented Architectures (SOA), and software-as-service, enable a shift from a program-centric IT to a process centric architecture, which can be referred to under the umbrella term services computing [13]. These technological developments allow for the transition of a service to a digital stratum, also known as IT services. IT services may comprise one or many business transactions and functions, and can be accessed without any knowledge of their underlying implementation details [14]. These IT services are reusable objects that embody repeatable business activities and tasks, and can be accessed through a network [4]. Despite technological developments in service oriented technologies, there is still a lack of understanding of how these IT resources should be leveraged, modified, and examined in order to create value and result in a competitive advantage

With the encapsulation of services into IT fragments, we argue that the service science paradigm should be adapted accordingly. The fundamental process of deploying services of value, as viewed by service science, is a result of operant resources. Accordingly, the operant resources under the information technology spectrum should be outlined. Operant resources under this category should capture the processes by which IT can enable change and result in new service offerings. Additionally, these processes should not be restricted to actions within the firm, but should also take into account the changing business environment as well as cooperation with the extended value network. In the following section we introduce the DCs view which we deem a pertinent theoretical foundation on which we can actualize our rationale in order to define operant resources that IT should enable. Dynamic capabilities are considered as being an equivalent subset of operant resources on services, and are therefore necessary in order to deliver value propositions in service-oriented enterprises [15].

3 Dynamic Capabilities

Within the domain of information systems, a number of theories from other disciplines such as economics, computer science, psychology, and general management have been adapted and applied [5]. The Resource Based View (RBV) of the firm has been the predominant theoretical lens of examination of a firms IT resources and capabilities for the past two decades [16]. According to the RBV, resources which are Valuable, Rare, Inimitable, and Non-Substitutable (VRIN-ness), provide firms with a competitive edge when applied in a suitable context. Researchers and practitioners that have used the RBV, refer to IT using various terms, such as resources, competencies, skills, and capabilities. Each of these terms captures a complementary aspect of IT. However, the RBV does not provide any theoretical basis on how a firm fosters change when the business environment shifts [17]. Hence, the RBV may not be the most appropriate theoretical lens to examine IT, since it does not address how future valuable resources can be created, or how the current stock of VRIN resources can be refreshed in turbulent conditions [18].

Dynamic Capabilities (DC) are the firms' capacities to integrate, build, and reconfigure internal and external resources/competencies to address and shape rapidly changing business environments [19]. In a knowledge-based, intangible service-dominant and quasi-globalized economy, characterized by fierce competition, uncertainty and change, it is DCs that are more important for the effectuation of sustainable competitive advantage [20]. To a greater extent than others, service-oriented enterprises rely on being able to develop competences such as skills and knowledge in order to constantly enhance their service provisioning [21]. At this point it is important to provide a distinction between dynamic capabilities and operational capabilities. While the former describe the ability to reconfigure and change the way a firm operates, the latter imply the ability to "make a daily living" [22]. Operational capabilities can be applied to execute routine activities; however, they do not address the processes of change. DCs are about developing the most adequate resource base for the near or distant future, reflecting thus, the process that impacts upon resources. Consequently, the mechanism of change includes the interaction of dynamic capabilities with the resource base. We argue that service firms operate in such dynamic environments; consequently a suitable theoretical prism to examine them would be under the DCs view.

A great deal of debate has emerged regarding what can be considered as a dynamic capability. According to [23], capabilities can be distinguished into four typologies. The first are those that denote an ability to perform basic activities. These types of capabilities are identified through the RBV of the firm, and with respect to IT capabilities have been detailed in literature [5]. The second category concerns capabilities that effectuate a dynamic improvement to resources and activities of a firm. The third category resemble the second since they both address dynamic improvements, but differentiate in that they are about developing strategies and resources before competitors. The second and third categories of capabilities are in congruence with the definition of dynamic capabilities by Teece et. al. [19]. The fourth category is about higher order or meta-capabilities, which try to capture how a firm learns to learn capabilities.

Apart from the definitions and distinctions of DCs in relation to other concepts, it has not been until recently that a growing body of research has attempted to analyze and measure them. The most exercised approach in quantifying DCs is by building upon [24] (reconfiguring, learning, integrating, and coordinating) and [19] (sensing the environment to seize opportunities and reconfigure) definitions. This view of dynamic capabilities perceives them as comprising of processes concerned with sensing, learning, integrating, and coordinating activities. Through this view, DCs are proposed as a tool in order to renew existing operational capabilities [25]. In a similar manner, some articles see DCs as the sensing, seizing, and transforming capacity, thus providing a slightly differentiated view [26]. Studies that have applied these two conceptualizations of DCs, examine their impact either through an aggregate manner, or in isolation.

In the following section we build upon the consideration aforementioned about what DCs include and how they should be assessed, and propose a set of enabling IT capabilities. We commence the conceptualization by identifying the DCs that are of essence to service-oriented enterprises in turbulent business environments, thus applying an expected-outcome view. We do so by reviewing MIS and service science literature and ground our taxonomy on past DCs empirical studies. In sequence, we explain their interdependencies, and delineate the mechanisms and conditions through which they enable change in operational capabilities.

4 IT-Enabled Dynamic Capabilities

Dynamic capabilities have been to date restricted to capacities of the business function, perceiving IT only as a supportive resource. IT has mostly been conceptualized in terms of a resource and has long been documented and examined in literature. Very few studies however examine IT through its enabling impacts on effectuating change, and the capabilities which are required to do so which are increasingly been regarded in literature as the source of competitive performance. Simply put, recent IS literature posits that the mere presence of IT is not sufficient to lead to a competitive advantage, but what are of importance are the capabilities that accrue from the effective competence to leverage it [27]. A static view of IT resources may be suited in relatively stable environments, however, in dynamic environments, what is of essence is examining the enabling capability of IT for competitive survival and adaptation. The means by which firms` resources are renewed, can be examined through the lens of IT-enabled dynamic capabilities. This view is particularly relevant since the role of IT have evolved from a tool to support business processes, to become an integral fabric of the enterprise in which business and IT are fused [27].

We commence by identifying IT-enabled dynamic capabilities by drawing on literature within the DC domain that conceptualizes their dimensions. The proposed dynamic capabilities are reinforced by studies within the Information Systems (IS) domain that stress the importance that firms must place in developing them. In order to distinguish dynamic capabilities we follow the suggestion of Fink & Neuman [28] who note: *"Dynamic capabilities actually consist of identifiable and specific routines that often have been a subject of extensive empirical research in their own rights"*. The ultimate goal of dynamic capabilities is to enable evolutionary fitness and prevent rigidities in turbulent environments [19]. There is a general consensus in research that the subset of dynamic capabilities that enable this adaptive capability can be distinguished into: sensing, learning, integration, and reconfiguration capabilities [24], [25]. Although these proposed IT-enabled capabilities are deemed as important enablers for surviving in a competitive landscape, they are not exhaustive. This means that there may be other IT-enabled dynamic capabilities that are of importance in certain contexts.

4.1 IT-Enabled Coordination/Reconfiguration

The capacity to respond to changing business needs is of paramount importance in today's enterprises, especially those operating in the service industry. The ability to effectively coordinate resources to enable fit with the environment has been referred to in business and IT literature by a number of different terms such as agility [29], flexibility [30], and adaptiveness [31]. All of the aforementioned concepts address the same underlying need of enterprises, to be capable of change when the external environment requires it. This notion of change connotes that it must be performed in a timely manner in order to be effective.

When a firm deems it necessary to reposition itself in the changing environment, it must reconfigure many resources and relationships within and across its boundaries to implement these decisions [32]. The notion of adaptation to the changing environment entails the coordination and reconfiguration of existing resources rather than the development of novel ones which will fundamentally reshape it. The difference in these two states is that in the former, a firm senses the need to reposition itself and renews resources accordingly, while the latter implies an exploration perspective into novel solutions, thus revolutionizing the domain in which it operates.

In service-oriented enterprises, resources are heavily dependent on IT and in most cases even fused in the form of IT services. Recent publications in the area of MIS, have stressed the importance of examining the co-evolution of *physical* and *social* technologies [7]. According to Nelson`s theorization, *Physical Technologies* include everything we refer to as Information and Communication Technologies (ICTs), while *Social Technologies* are concerned with organizing work and people, i.e. the information systems competencies and management [33]. The vast majority of studies have been concerned with examining the degree to which physical technologies can be subjected to change [34]. This capacity is known as IT flexibility. The development of service-oriented technologies was propelled by the need for more flexibility with regard to physical technologies; however, the social technologies required in order to reconfigure them are still inadequately examined [29]. Although research has highlighted the importance of possessing flexible IT resources, there is limited attention on its ability to mobilize and coordinate IT resources to adapt to changing environments. This capability becomes particularly important when business logic is fused in IT in the form of digitized business processes. We therefore perceive IT-enabled coordination/reconfiguration as an important dynamic capability.

The scale, scope, and frequency of reconfiguration of IT becomes greater as environmental turbulence increases. Therefore, the capability of doing so in a timely and effective manner constitutes it as a dynamic resource of the firm. We define *IT-enabled coordination/reconfiguration* as the capacity of a firm to rapidly reconstruct its internal and external IT services in order to respond to the changing business environment. This definition encompasses the ability of co-evolution at a physical and social technology level simultaneously. It is evident from the definition that the capacity of IT-enabled coordination/reconfiguration entails the notion of dynamic alignment, i.e. to be capable of evolving business and IT strategy concurrently.

4.2 IT-Enabled Integration

A central consideration of service science is to strengthen ties between business collaborators, customers, and other value co-creators. It is through the interconnectedness and synergies of these service systems that value is realized. The value of collaboration between firms has been noted in many scholarly publications and in particular in the area of strategic alliances [35]. The need to engage in such collaborative schemes is usually propelled by the competitive landscape in which they operate [36].

The idea of IT-enabled collaboration has been mostly concerned with how to streamline operations of the supply value chain with the aim of reducing costs, human error, and time-to-completion. This concept is noted as IT integration in academic literature, and has been subject to extensive research regarding the value derived. An alternative term is Enterprise Integration (EI) which goes beyond systems interoperability, to include information flows, and knowledge sharing among firms [37].

The notion of moving beyond the limits of the traditional value chain on to what is known as the extended network of enterprises is considered as the next step in IT integration and EI [7]. There is a growing notion in service science to move beyond traditional supply chains that provide integration between suppliers and customers to what are called value networks. Value networks are spontaneously sensing and spatially responding temporal structures of loosely coupled social and economic actors that aim to co-produce service offerings, exchange service offerings, and co-create value [38]. The benefits from such arrangements are not limited to business process automation, and cost reduction [39], but also include an increased capacity of intra-preneurship [40]. The emerging service-oriented technologies enable firms to connect to complex networks of collaborators in a flexible manner. This network paradigm is facilitated through service-technologies that alleviate past technical constraints and allow firms to loosely engage and disengage in value networks. However, the exploitation of distributed resources and the ability to rapidly connect to them cannot be performed without a cultivated and mature integration capability.

Hence, we propose the notion of IT-enabled integration. We define the term as the capacity of a firm to rapidly embed IT with current or new nodes of the extended network in order to obtain access to remote resources that are of significance. Despite IT integration being addressed extensively in literature, we view it from a different standpoint. While past studies examine integration as the degree to which connections are established, we differentiate IT-enabled integration as the capability and agility to incorporate dispersed resources in a firm's value network. We consider that this capability encompasses the competence of integrating resources, such as knowledge and operations to the firm's asset base, but does not include the know-how of leveraging them.

4.3 IT-Enabled Learning

The importance of embedding new knowledge in service offering is at the cornerstone of service science. One of the main reasons firms engage in co-creation is so that they can leverage unique resources and knowledge that co-creators possess. Much has been written about how organizations learn and incorporate knowledge into their operations. The ability of firms to do so is noted as a critical antecedent of competitive advantage. A number of research streams such as, organizational learning, knowledge-view of the organizations, and absorptive capacity emphasize on the importance of learning at the firm level, and attempt to unravel the processes through which knowledge can be leveraged as a strategic asset [41]. Since the service-based economy is predominantly fueled by knowledge, the primary source of advantage is through the effective accumulation and exploitation of such intangible assets. By effectively utilizing knowledge resources, firms can propel their innovation performance and enhance their operational efficiency. This fact is noted by Cohen and

Levinthal [42] whom suggest that learning at the enterprise level of analysis can render them as more proactive rather than reactive.

Although IT-enabled integration capability facilitates firms to extend their locus of reach and develop boundary-spanning capabilities [43]; the ability to identify critical resources and assimilate new knowledge and competencies into operations are part of a firm's IT-enabled learning. We build on absorptive capacity theory and define IT-enabled learning as the capability of a firm to leverage IT in order to acquire, assimilate, transform, and exploit knowledge [44]. Although there has been much attention on the potential of IT resources in aiding knowledge management at the firm level, the predominant view of assessing the impact of IT has been to examine if certain systems have been deployed. The notion of IT-enabled learning does not center on the presence or absence of a particular IS, but includes aspects of competency in leveraging the full potential of IT for such tasks.

4.4 IT-Enabled Sensing

A basic capacity for any firms` competitive survival is to be competent in repositioning itself in the changing business environment. Even if firms are very effective is reconfiguring their operational capabilities and service offerings, if they are not alignment with market demands they are doomed to failure. Therefore, the surveillance of market trends is a critical in order to ignite an appropriate reaction. A sensing capability is especially important in high volatility environments such as the ones service-oriented enterprises operate in. The continuous shift of market requirements, the rapid technological development rates, and the fluid socio-economic circumstances in the global plateau make the positioning of firm difficult if it does not have the appropriate mechanisms to decipher the external environment. Hence, it is imperative to invest in capabilities that can precisely evaluate and detect shifting conditions.

The importance of developing IT capabilities that enable firms to stay responsive to market intelligence by affectively sensing the environment is argued in literature [27]. We define IT-enabled capability as the capacity of a firm to effectively utilize IT in order to generate, disseminate, and respond to market intelligence [19]. The generation of marketing intelligence requires close collaboration with customers, and being able to detect competitors' moves. Dissemination is revolved around making sense of events and exploring new opportunities. Finally, responding to market intelligence includes formulating plans and pursuing specific market segments. IT used to support these functions can vary significantly, from social media platforms, to web services, market analytics software etc. Since firms may chose varying IT resources to enable their sensing capability, and since deploying such systems doesn't automatically mean an enhanced capacity to scan the environment, we argue that it is best to focus on the overall IT-enabled capability of doing so [45]. Despite there being equifinal solutions in achieving IT-enabled sensing capability, this does not mean that some IT resources are more appropriate than others in certain contexts. However, in order to examine the effect of IT-enabled sensing on competitive advantage, it is more appropriate to do so by evaluating the companies leveraging competence.

5 Discussion and Implications

In the present study we apply service science logic on future trajectories as is suggested in MIS literature. The aim is to define the processes by which value is created in the fast paced, quasi-globalized, and turbulent service economy, as a result of the co-evolution of physical and social technologies. We specifically focus on service-oriented enterprises since they are in alignment with the logic proposed in service science. In such firms, the predominant way of thinking about value is in terms of services, which are increasingly dependent upon IT. According to service science, competitive advantage is a result of operant resources that facilitate the development and renewal of service offering. To this end we analyze MIS literature and survey the future quests that IT must address. We conclude that IT has an enabling role by enabling renewal of resources and competencies. In order to highlight the role of IT in enabling competitive advantage, we employ a dynamic capabilities perspective and propose IT-enabled dynamic capabilities that firms must cultivate. Through this way we manage to define the respective operant resources of service science, as dynamic capabilities for the co-evolution of physical and social technologies. We arrive at four dynamic IT resources that are critical in order to compete in turbulent environments. These are: *IT-enabled Coordination/Reconfiguration, IT-enabled Integration, IT-enabled Learning, and IT-enabled sensing*. The term dynamic in these definition denotes their capacity to incur change in operational capabilities and resources.

The main contribution of this research is that it combines the principles of service science with open questions of the MIS literature. Thus, we put forth an example of operationalizing the service science paradigm by including the DC theory to explain the aspect of change (operant resources). The main differentiation of the current study in comparison with others is that we develop IT-enabled dynamic capabilities that capture the interwoven nature of business and IT in service-oriented enterprises. Therefore, we create a new perspective of viewing IT capabilities.

Despite the novelty of the approach and the capabilities identified, our research has certain limitations. First and foremost we have yet to determine how these IT-enabled capabilities are developed which is the essence of the DCs view. Our study has been limited in describing the necessary DCs. What remains is to determine what resources, skills, competencies, and organizational structures enable their formation. To actualize this objective, a structured literature review is required, in which the antecedents that support these IT-enabled capabilities can be outlined. This will enable the formulation of a clear and coherent roadmap for managers to develop IT-enabled capabilities that are necessary for their survival. In literature, the delineation of dynamic capabilities is examined in two ways, by defining the interplay between other higher order dynamic capabilities, or by identifying the antecedents in terms of structures, and conditions; also known as micro-foundations. A second consideration that remains unaddressed is under what circumstances and through what mechanisms these dynamic resources impact a firm's competitive position. Past research on DCs indicates that value can only be derived in conditions that invoke the use of them, such as highly turbulent and changing environments. This means that a firm that develops dynamic capabilities and operates in a stable environment may not necessarily perceive performance gains. Therefore, any future studies attempting to identify the outcomes of IT-enabled dynamic capabilities should consider this fact. Additionally, the effects of these particular capabilities may not be directly linked to sustained competitive advantage, but could be mediated with relevant concepts such as innovation capability or business agility. This is a particularly interesting aspect since through this novel perspective past shortcoming of IS strategy literature may be explained.

References

1. Spohrer, J., Maglio, P.P.: The Emergence of Service Science: Toward Systematic Service Innovations to Accelerate Co-Creation of Value. Production and operations management 17, 238–246 (2008)
2. Berry, L.L., Shankar, V., Parish, J.T., Cadwallader, S., Dotzel, T.: Creating new markets through service innovation, Cambridge (2006)
3. Vargo, S.L., Maglio, P.P., Akaka, M.A.: On value and value co-creation: A service systems and service logic perspective. European Management Journal 26, 145–152 (2008)
4. Demirkan, H., Kauffman, R.J., Vayghan, J.A., Fill, H.-G., Karagiannis, D., Maglio, P.P.: Service-oriented technology and management: Perspectives on research and practice for the coming decade. Electronic Commerce Research and Applications 7, 356–376 (2009)
5. Wade, M., Hulland, J.: Review: The resource-based view and information systems research: Review, extension, and suggestions for future research. MIS Quarterly 28, 107–142 (2004)
6. Ambrosini, V., Bowman, C.: What are dynamic capabilities and are they a useful construct in strategic management? International Journal of Management Reviews 11, 29–49 (2009)
7. Merali, Y., Papadopoulos, T., Nadkarni, T.: Information systems strategy: Past, present, future? The Journal of Strategic Information Systems 21, 125–153 (2012)
8. Barile, S., Polese, F.: Smart service systems and viable service systems: Applying systems theory to service science. Service Science 2, 21–40 (2010)
9. Maglio, P.P., Spohrer, J.: Fundamentals of service science. Journal of the Academy of Marketing Science 36, 18–20 (2008)
10. Madhavaram, S., Hunt, S.D.: The service-dominant logic and a hierarchy of operant resources: developing masterful operant resources and implications for marketing strategy. Journal of the Academy of Marketing Science 36, 67–82 (2008)
11. Mikalef, P., Pateli, A.: A systematic meta-analytic review on factors influencing the strategic alignment in Service-Oriented Architecture projects. In: Proceedings of the 6th Mediterranean Conference on Information Systems, MCIS, Limassol, Cyprus (2011)
12. Zhao, J.L., Tanniru, M., Zhang, L.-J.: Services computing as the foundation of enterprise agility: Overview of recent advances and introduction to the special issue. Information Systems Frontiers 9, 1–8 (2007)
13. MacKenzie, C.M., Laskey, K., McCabe, F., Brown, P.F., Metz, R., Hamilton, B.A.: Reference model for service oriented architecture 1.0. OASIS Standard 12 (2006)
14. Vitharana, P., Bhaskaran, K., Jain, H., Wang, H., Zhao, L.: Service-oriented enterprises and architectures: state of the art and research opportunities (2007)

15. Arnould, E.J., Price, L.L., Malshe, A.: Toward a cultural resource-based theory of the customer. The service-dominant logic of marketing: Dialog, debate and directions, pp. 320–333 (2006)
16. Wernerfelt, B.: A resource-based view of the firm. Strategic Management Journal 5, 171–180 (1984)
17. Teece, D., Pisano, G.: The dynamic capabilities of firms: an introduction. Industrial and Corporate Change 3, 537–556 (1994)
18. Ambrosini, V., Bowman, C.: What are dynamic capabilities and are they a useful construct in strategic management? International Journal of Management Reviews 11, 29–49 (2009)
19. Teece, D.J.: Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal 28, 1319–1350 (2007)
20. Pitelis, C.N., Teece, D.J.: The (new) nature and essence of the firm. European Management Review 6, 5–15 (2009)
21. Lusch, R.F., Vargo, S.L., O'Brien, M.: Competing through service: insights from service-dominant logic. Journal of Retailing 83, 5–18 (2007)
22. Winter, S.G.: Understanding dynamic capabilities. Strategic Management Journal 24, 991–995 (2003)
23. Collis, D.J.: Research note: how valuable are organizational capabilities? Strategic Management Journal 15, 143–152 (1994)
24. Teece, D.J., Pisano, G., Shuen, A.: Dynamic Capabilities and Strategic Management. Strategic Management Journal 18, 509–533 (1997)
25. Pavlou, P.A., El Sawy, O.A.: Understanding the elusive black box of dynamic capabilities. Decision Sciences 42, 239–273 (2011)
26. Lichtenthaler, U., Ernst, H.: The performance implications of dynamic capabilities: The case of product innovation. Journal of Product Innovation Management (2012)
27. El Sawy, O.A., Pavlou, P.A.: IT-Enabled Business Capabilities for Turbulent Environments. MIS Quarterly Executive 7 (2008)
28. Fink, L., Neumann, S.: Exploring the perceived business value of the flexibility enabled by information technology infrastructure. Information & Management 46, 90–99 (2009)
29. Tallon, P.P.: Inside the adaptive enterprise: an information technology capabilities perspective on business process agility. Information Technology and Management 9, 21–36 (2008)
30. Power, B., Reid, G.C.: Flexibility, firm-specific turbulence and the performance of the long-lived small firm. Review of Industrial Organization 26, 415–443 (2005)
31. Tallon, P.P., Pinsonneault, A.: Competing Perspectives on the Link Between Strategic Information Technology Alignment and Organizational Agility: Insights from a Mediation Model. Mis Quarterly 35 (2011)
32. Tanriverdi, H., Rai, A., Venkatraman, N.: Research commentary-Reframing the dominant quests of information systems strategy research for complex adaptive business systems. Information Systems Research 21, 822–834 (2010)
33. Nelson, R.R.: Physical and social technologies and their evolution. Économie Appliquée 56, 13–32 (2003)
34. Byrd, T.A., Turner, D.E.: Measuring the flexibility of information technology infrastructure: Exploratory analysis of a construct. Journal of Management Information Systems 17, 167–208 (2000)
35. Das, T.K., Teng, B.-S.: A resource-based theory of strategic alliances. Journal of Management 26, 31–61 (2000)

36. Eisenhardt, K.M., Schoonhoven, C.B.: Resource-based view of strategic alliance formation: Strategic and social effects in entrepreneurial firms. Organization Science 7, 136–150 (1996)
37. Vernadat, F.B.: Enterprise modelling and integration. Springer (2003)
38. Lusch, R.F., Vargo, S.L., Tanniru, M.: Service, value networks and learning. Journal of the Academy of Marketing Science 38, 19–31 (2010)
39. Tanriverdi, H., Uysal, V.B.: Cross-business information technology integration and acquirer value creation in corporate mergers and acquisitions. Information Systems Research 22, 703–720 (2011)
40. Benitez-Amado, J., Llorens-Montes, F.J., Perez-Arostegui, M.N.: Information technology-enabled intrapreneurship culture and firm performance. Industrial Management & Data Systems 110, 550–566 (2010)
41. Lichtenthaler, U.: Absorptive capacity, environmental turbulence, and the complementarity of organizational learning processes. Academy of Management Journal 52, 822–846 (2009)
42. Cohen, W.M., Levinthal, D.A.: Absorptive capacity: a new perspective on learning and innovation. Administrative Science Quarterly 35 (1990)
43. Rosenkopf, L., Nerkar, A.: Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. Strategic Management Journal 22, 287–306 (2001)
44. Zahra, S.A., George, G.: Absorptive capacity: A review, reconceptualization, and extension. Academy of Management Review 27, 185–203 (2002)
45. Mikalef, P., Pateli, A., Batenburg, R., van de Wetering, R.: Business alignment in the procurement domain: a study of antecedents and determinants of supply chain performance. International Journal of Information Systems and Project Management 2, 43–59 (2014)

Open Sustainability Innovation—A Pragmatic Standpoint of Sustainable HCI

Moyen Mohammad Mustaquim and Tobias Nyström

Department of Informatics and Media, Uppsala University,
Box 513, 751 20 Uppsala, Sweden
{moyen.mustaquim,tobias.nystrom}@im.uu.se

Abstract. The importance of sustainability in design became a major topic of interest in HCI research. Past research has shown how classical HCI design principles could be used to create values of the design through the open innovation concept. The scope of this paper follows from an earlier work of open innovation design principles that established the basis of open sustainability innovation since contributions from a dissimilar applied form of HCI could be promising. We in this paper have analyzed the concept of open sustainability innovation from the perspective of sustainable HCI. The paper outlines seven design principles for open sustainability innovation and has illustrated them in the form of a framework. The notion of open sustainability innovation and sustainable HCI are then revisited and the role of sustainable HCI for developing sustainable products, services, and initiatives using open innovation are clarified.

Keywords: Open innovation, Open Sustainability Innovation, Sustainable HCI, Design Principles.

1 Introduction

Sometimes referred to as a creative destruction activity, organizations are often considered by many to be the main causes that create a lack of sustainability in the society [21]. Even though it is a relatively new concept and approach for innovation, open innovation is considered to be a major shift for many organizations since often organizations could be restricted by their limited knowledge and resources. At the same time, very little attention has so far been given to the open innovation concept to understand its role for shaping and building a sustainable future. Perhaps this idea could play a great role in promoting sustainability if innovation strategy could be altered with a focused desired goal. Indeed the concept of open sustainability innovation was elevated when it was used to develop ideas or initiatives, products or services that were sustainable. At times this approach is practiced by organizations in their marketing phase for introducing new information to their consumers, given that the customers were totally unaware of a new product or service [1]. Sustainable HCI is a growing area of research concerning everyday practices focusing on sustainability, despite the fact that its scope should not only be limited to everyday

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 101–112, 2014.

life but instead the approach towards sustainability for emerging in everyday life could be thought as a basis on which it is built. [18] Sustainable HCI research and studies (for example [12, 13]) contributed to supporting sustainable decision-making for product-purchasing [20] and on the whole it is important to understand the behavior of the consumer if HCI research would like to provide more support for this. We in this paper used the notion of sustainable HCI to enhance the open innovation concept towards open sustainability innovation. The underlying research question was: “How could sustainable HCI create values in marketing for developing sustainable products, services, and initiatives by using open innovation?” The structure of this paper’s argument was illustrated in the form of a block diagram shown in Figure 1. So far, previous research has shown how open innovation design principles could be formulated from HCI design principles. We took a similar approach and used HCI design principles to add values to the domain of sustainable HCI. We then used the enhanced concept of sustainable HCI as an approach for constructing values in marketing, aiming for sustainability goals to be elevated by using open innovation. The paper presented seven design principles for open sustainability innovation and a framework was proposed to explain how these design principles would work for promoting sustainability through open innovation. Finally we revisited the notion of sustainable HCI from the context of open sustainability innovation and redefined sustainable HCI in that context, in which it was argued that sustainable HCI could be more than creating products or services to change the consumer’s behavior towards any ecological action.

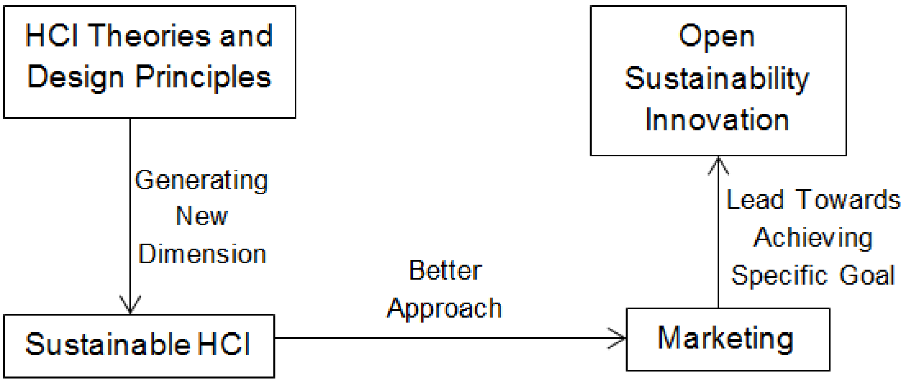


Fig. 1. Sustainable HCI and its impact towards open sustainability innovation

2 Background

To grasp the connection between open sustainability innovation and sustainable HCI, it was important to explore some theoretical foundations. In section 2.1 the background for open sustainability innovation is presented, while section 2.2 discusses sustainable HCI in connection to open innovation.

2.1 Open Innovation and Innovation for Sustainability

By many, open innovation is considered to be a radical change in setting innovation strategy and managing innovation. Introduced a decade ago by Henry Chesbrough, the open innovation concept is debated as being something old to be packaged in a new format [22]. He stated that not any longer can a single organization have all the resources or knowledge that is needed for creating innovation themselves [4] and therefore they need more inputs from outside the organization's boundary. Previous research revealed that a collective intelligence as an approach for creation of ideas could often be better than a single mind. [2] Open innovation is a concept in which more stakeholders get involved in the innovation process compared to its opposite—"closed innovation", when the company innovates in solitude and behind closed doors. The stakeholders in the case of open innovation could, for instance, be suppliers, customers, or competitors.

In addition, no general definition of sustainability exists since it is dependent on the context and the research field (for comprehensive reviews of sustainability terms see [10]). Nevertheless, the negative impact that human activities could have on our environment is acknowledged in sustainability. One key definition for sustainable development was stated in the Bruntland Comission: "… development that meets the needs of the present without compromising the ability of future generations to meet their own needs." [23] The Bruntland Commission opened the awareness that existing processes for development, consumption, and production could be unsustainable. Thus sustainability would be about improving our life, not creating a negative impact but instead minimizing, if possible reversing or reducing the negative impacts. At the same time, observed sustainability research in information technology is mainly focused on technical solutions to reduce the emission of carbon dioxide from IT hardware [5].

Charter et al. [3] stated: "Marketing has a key strategic role as the interface between consumption and production, with a considerable influence over the construction of the company's product/market portfolio and also over its communications efforts." Environmental performances together with social indicators were seen as a key service and product attribute. This could be a source for differentiation to give potential competitive advantage [19], and thus both the innovation as well as the marketing process is important as a bundle to be offered in the market. Again, to achieve competitive advantages the marketing policy must be in alignment with the innovation strategy. Peattie and Peattie [17] pointed out the ineffectiveness of green marketing by claiming: 'Creating meaningful progress towards sustainability requires more radical solutions than just the development of new products and product substitutions amongst consumers.' A better solution to this problem could be to get more stakeholders involved in the innovation and marketing process. Open sustainability innovation uses the open innovation concept with an aim to reach sustainability through the marketing mechanism. With more stakeholders involved in the innovation process, chances of reaching sustainable goals could increase and this should be applicable for marketing too. Active stakeholder involvement is central for successful use of open innovation, as without it no new ideas outside the organization would be fetched.

2.2 Sustainable HCI and Open Sustainability Innovation

Research with a component of sustainability is ubiquitous and sustainability research in HCI is accelerating. The growth volume of academic papers about sustainable HCI has increased [8] and is now one of the fastest growing areas in HCI [9]. By reading a review made by DiSalvo, Sengers and Brynjarsdóttir [7] in 2010, it seems that there is a widespread use of sustainable HCI in research. Furthermore, the focus of sustainability research in IT [5] mimics the research in sustainable HCI, which is focused on the reduction of resource use and minimization of the emission of carbon dioxide through the design of systems that might change individuals' consumption and choice behavior [9]. The focus on shaping the individual's choice based on a negative motivation in order to reduce consumption is therefore a prevalent theme for research in sustainable HCI [9]. One of the two suggested promising directions for future research in which HCI would be related to the environmental issue was "participatory design" [11], when the issues of participatory design could be answered by including stakeholders in the innovation and marketing processes. In fact, sustainable system design should have a better chance to reach sustainable goals with more stakeholders involved, e.g. open innovation [15]. Thus an improved open innovation and marketing policy could promote positive incentives towards reaching different sustainable goals. In short, engaging users in the system development procedure with a goal of promoting sustainable action could be seen as a merger between sustainable HCI and open sustainability innovation research.

3 Open Sustainability Innovation Design Principles

Discussions from the background section clarified that open sustainability innovation initiative could take place as an enhanced approach of marketing by the organizations. Our aim in particular would be to use HCI design principles for adding new values to a marketing policy that organizations previously were unaware of, and to finally trigger the process of open sustainability innovation. At the same time, by using HCI design principles, the traditional understanding of sustainable HCI could possibly be taken into a new dimension and a new research process could be initiated.

We considered the "Buying Decision Process" approach by Kotler [14] as the principle basis of our research. Although first published by John Dewey [6] in 1910, the elements of this approach have been reformed over the years [20]. Moreover we took into account the characteristics of sustainability marketing, addressed by Belz et al. [1] as our second basis for formulating design principles. According to Belz et al. these characteristics are important to understand for organizations to build and maintain sustainable relationships with their customers. For our HCI design principles we chose the universal design approach and its principles. Mustaquim and Nyström [16] showed how open innovation design principles could be shaped from the angle of universal design. In their research it was shown how universal design could create values in other domains (open innovation in this case) which was not limited within the scope of physical disabilities. [16] To put it differently, we used a similar concept of universal design and mapped design principles from HCI research domain into the two marketing approaches—one for creating better marketing strategies and the other for creating and maintaining better relationships with the clients.

Figure 2 shows a framework in which universal design principles are diagrammed with the five stage framework of the buying decision process model. In Figure 3 (see below) we have shown how the characteristics of sustainability marketing could be mapped with universal design principles.

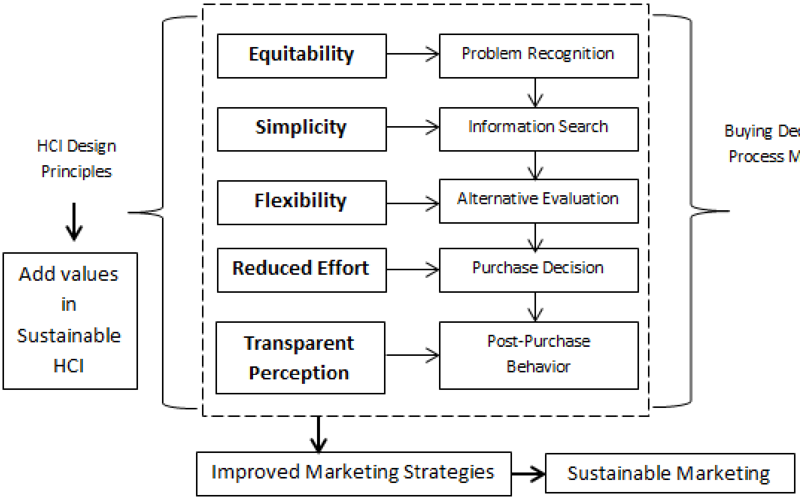


Fig. 2. A refined version of buying decision process [1] using HCI design principles

These two frameworks were formulated using a combinatorial approach in which two different theories of marketing and HCI design principles were considered. Nevertheless, both frameworks are aimed towards achieving sustainable marketing.

It was illustrated in Figure 2 how our approach could improve marketing strategies, which would lead towards sustainable marketing. The phase of problem-recognition could be equal in the sense that it should be well balanced between the customer's requirement and their desired state, while they would decide to purchase a product. The customers would then try to find information from different sources about the product they would intend to purchase and this process should be simple. The third phase would be evaluation of various alternatives that should be flexible to let customers evaluate the primary requirements, initiated in an equitable state of problem-recognition. However, customers might often not follow these steps to come up with a purchase decision. They could just skip these steps and buy a product without collecting too much information about it. Hence we added the reduced effort design principle with the purchase decision factor. Subsequently, information about a product could be very powerful [20] for making customers buy the product and it should be designed by keeping in mind that too much effort should not be required to spend for understanding the information about it. In fact this would make the direct purchase decision process easier for the customers. The final phase from the buying decision process model was post-purchase behavior, mapped with the transparent perception design principle. When customers buy a product they might either like or dislike it; they could either recommend it to others or not. Customer decision on an action would depend on how a product was designed and what effect it would leave on shaping the customer's attitude. In particular, if information provided by the designer were transparent and easy to perceive by the customers, it should create a positive post-purchase behavior for them.

Figure 3 took into consideration the characteristics of sustainable marketing and four design principles from HCI. If we want to maintain a product's cost parallel to the purchase cost, cost of use together with the post-use costs, a design should have a smooth approachability characteristic for it. Without being able to approach easily towards any of the cost properties, it would be a challenge for the organization to maintain the customer cost in general.

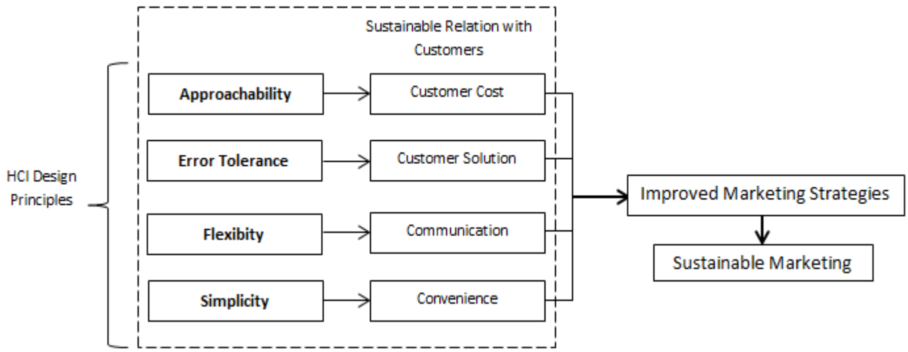


Fig. 3. Sustainable relationship with customers and HCI design principles

Further, the elements of customer cost are a complex concept that is not limited within the three types of costs mentioned earlier but could include phenomena like risks and qualities. Therefore easy approachability in design to have a smooth transaction between these different variables would be important for balancing the customer cost. Providing better customer solution is the next characteristic, which should maintain higher error tolerance. The concept of providing customer solution aiming to offer solutions for different issues when customers use a product is not enough in today's business world. For many, the customer solution could mean many other different things that could leave impacts on the customer's decision whether to buy products from the same organization in the future. Therefore error-free or higher error tolerance is important while designing customer solutions. Communication is the third character of sustainable marketing and we assigned flexibility together with this character. Better communication is always good for customers, but flexible communication between company and client could have a long-term impact on the purchase decision of a product from an organization. Finally, convenience was one of the key properties that could make customers choose to buy one product from a particular company. Convenience could create values in the customer's mind, which would be not just limited to the price. There could be several factors for producing success in convenience and a successful relation of these could create values for business. Again these factors could depend on the nature of business and product. We therefore took the simplicity design principle to map the convenience characteristic. Simplicity in design could lead towards convenience by relating other factors associated with the type of business or product. The approach of using these marketing characteristics together with selected design principles from HCI could improve marketing strategies and contribute positively towards sustainable marketing. In Table 1 we have shown a characteristic matrix of HCI design principles with the buying decision process and sustainable marketing features together with the bases for formulating open sustainability innovation design principles.

Based on the above discussion of the buying decision process and characteristics of sustainable relation with customer together with the design principles of HCI, we came up with seven design principles for open sustainability innovation. The proposed design principles are described in Section 3.1 below.

Table 1. Characteristic matrix for open sustainability innovation

Buying Decision Process and Sustainable Marketing Characteristics	Corresponding HCI Design Principles	Open Sustainability Innovation Design Principle Properties
Recognizing the Problem	Equitability	Steadiness between requirement and desired state of customer
Searching the Right Information	Simplicity	Simple information presentation
Evaluating Possible Alternatives	Flexibility	Ability to evaluate various alternatives
Decision of Purchase	Reduced Effort	Better understanding of the product
Post-Purchase Behavior of Customers	Transparent Perception	Enhanced information provided by the designers
Customer Costs	Approachability	Even flow between different cost variables
Customer Solution	Error Tolerance	Reduced error for efficiency
Communication	Flexibility	Improved way of maintaining communication
Convenience	Simplicity	Keeping the product design simple and thereby add value on it

3.1 Proposed Open Sustainability Innovation Design Principles

Principle one: Reduce gap between customer's requirement and desire for improved problem identification

Having open sustainability innovation as a main goal, an important thing would be to design for a specific need and not to simply just design a product. If the real requirement of the customers cannot be realized, they might end up buying a product anyway, even though they would not find that it gave them higher usability results. On the other hand, if the desire is too high then customers might never be satisfied with what they buy, regardless of whether an offered product improved usability or not. The gap between requirement and desire should therefore be reduced and this is an important thing to remember while designing a product or system to initiate sustainable practice.

Principle two: Make information presentation about a product for the customer an easy task
Customers should be able to find information about a product without much hassle and designers should remember this too during product design. Both an immediate and long-term decision-making process before purchasing a product could depend on how well customers were able to find information about it. Designers should focus on this matter since a successful presentation of information could leave optimistic impressions when balancing the decision of the customer's requirements and their desire to purchase a product.

Principle three: Expand the evaluation of product towards different alternative and enhance better communication with customer
Information provided to customers about a product in an appropriate style is significant, but it is better if customers are able to evaluate different products before they buy them. Designers should therefore reflect on this factor during design. Providing ability in evaluation to customers would open the door for improved communication with customers too. By providing improved information presentation about a product and better evaluation ability in contrast could create positive impacts on reducing the gap between customer's requirement and desire.

Principle four: Provide transparent information to customers for positive impact on their post-purchase behavior
Post-purchase behavior could trigger many things and one of them is whether or not it is going to reduce the gap between desire and requirement, i.e. providing clear and easily perceivable information to the client is important. If a design lacks in providing proper information or even provides limited information, the customer's way of using that product could be effected in a negative way. This could leave impact on other possible future clients who might share negative feelings about the product and even the organization while discussing with an existing client through word of mouth. Designers should thus remember these facts during design.

Principle five: Ensure balanced flow between different variables of customer cost
Customers cost is a complex factor and designers could make sure how to have a balanced flow between different variables of customer costs. In cost maintenance or reduction, whatever the goal might be, the success would depend on the relation with many other associated variables that build up total customer cost. Easy approachability from one factor to another would therefore be a key issue for ensuring better customer cost and should therefore be considered by the designers.

Principle six: Provide error-free customer solutions for efficiency
Ensuring error-free customer solution is important and there is no doubt about that. But how this factor could influence the marketing of a product might not always be very easy to realize. A product cannot just be designed for achieving many different goals without the designers being able to provide better instructions for support. This should be considered by designers for increased efficiency through their design.

Principle seven: Add value to product through simplicity in design for achieving convenience
Convenience and value are two issues that work with each other in the context of marketing the success of a product. By keeping the design simple, numerous values could be added to it, which would increase the convenience of a product. Simplicity is therefore a key factor that should be considered by the designers, which could help them achieve many goals including convenience and sustainability. On the contrary, designers should realize simplicity very contextually with specific design problems.

4 Proposed Framework

As an illustration, the proposed design principles are shown in the form of a theoretical framework in Figure 4 (see below) for explaining how they could work for the successful practice of open sustainability innovation. This would be a complex process, which is the result of a buying decision process and characteristics of

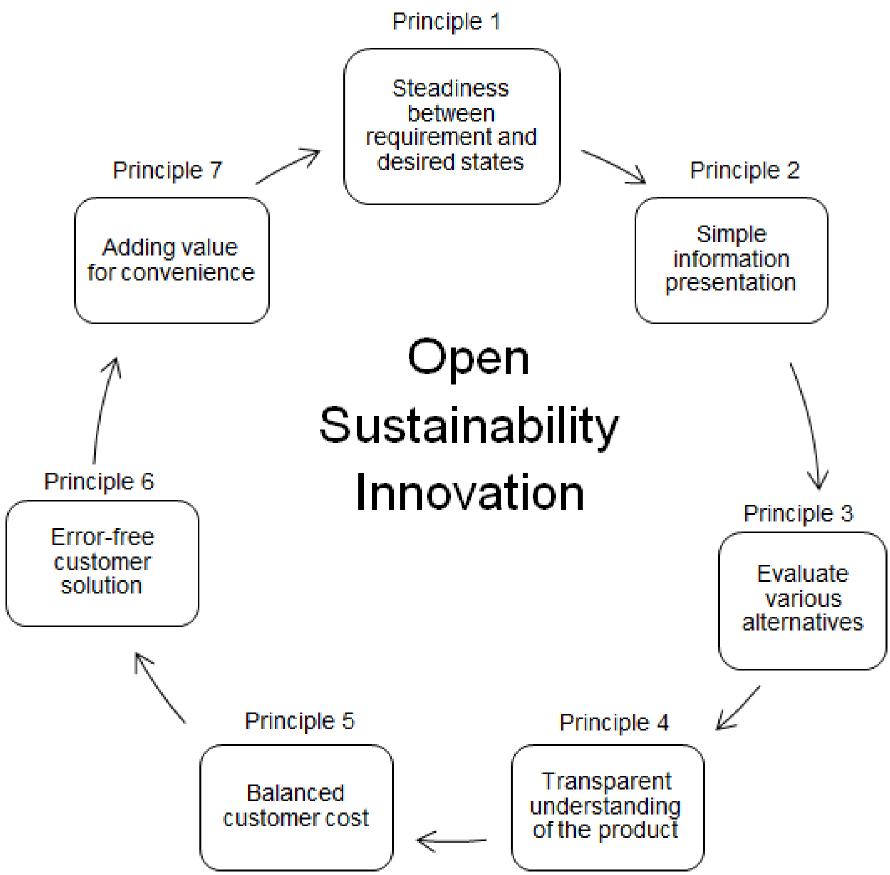


Fig. 4. A framework for open sustainability innovation process

sustainable marketing together with HCI design principles. The main goal here is to find a steady gap between the customer's requirement and the desire for the purchase of a product. Following the previous discussions of the design principles we could explore the proposed framework in Figure 4. Simple information presentation would allow customers to evaluate different alternatives before purchasing a product. This would follow successful transparent understanding of the product through the customer's post-purchase behavior. The customer cost and the variables associated with this could then be balanced, which would lead towards an error-free and efficient customer solution. These together could trigger the addition of positive values into a product for enhancing the convenience. Finally the cyclic process would achieve a steady balance between customer requirement and their desired state. If this could be achieved it would then be possible to claim that sustainability was the targeted achieved goal by the practice of the open innovation concept. Designers could be benefited from the use of this framework for enhancing a product development life cycle that could result in a sustainable end product. Besides, organizations practicing open innovation could consider this framework to make their innovation practice more sustainable and develop sustainable end products.

5 Discussions

The proposed framework was an abstraction of a very complex procedure that involved marketing, decision process, and HCI design principles. A few interesting research initiatives could be drawn from the outcome of this paper. First, the concept of sustainable HCI could be thought outside of the traditional box. Design principles from HCI could add value to the understanding of sustainable HCI. Second, we used universal design principles in this paper to compare the characteristics selected from two different theoretical bases. This showed us how universal design could be thought outside its traditional concept of accessibility issues, which is identified as a limitation. Previous research has argued and showed how a universal design concept could be expanded beyond the physical disability domain and this paper placed another milestone for adding values to this argument. Thus sustainability in HCI should not be limited only within the scope of creating persuasive technologies to change user behavior. Instead the notion of sustainable HCI could also mean the design process or design for achieving sustainability. Besides, adding different factors to lead a design procedure and for the end product to be sustainable could further be considered in the study of sustainable HCI. Open innovation is one such procedure that was explored in this paper. Open sustainability innovation design principles proposed in this paper could be empirically verified and the proposed framework could then be altered and improved as required. It would be interesting to see how an organization practicing open innovation could take the proposed design principles into practice and then to observe how sustainable their resulting end product would be. How sustainable the overall product development process could become, considering the proposed design principles, would be worth exploring too. The role of sustainable HCI in open sustainability innovation is therefore promising and further research could find interesting correlation between different fields of research as we did here with marketing and decision process with HCI. In contrast this could help to analyze the influence of other research fields in sustainability achievement.

6 Conclusions

Two different research disciplines were considered and compared with the design principles originated from HCI to formulate the design principles for open sustainability innovation. The design principles were then structured into a theoretical framework for explaining how they would be working in practice. It was shown that the sustainable HCI concept could be expanded from its traditional understanding and this could be done by the use of existing design principles from HCI. Besides we understood that the universal design concept could be thought outside of the accessibility domain and thus be considered as a factor to enhance usability in other design issues. A way of looking into the effect of sustainable HCI on open sustainability innovation process could thus be viable and positive for designing products and services with a sustainability goal.

References

1. Belz, F.-M., Peattie, K.: Sustainability Marketing. John Wiley & Sons Ltd., West Sussex (2012)
2. Bonabeau, E.: Decisions 2.0: The Power of Collective Intelligence. MIT Sloan Management Review 50(2), 45–52 (2009)
3. Charter, M., Peattie, K., Ottman, J., Polonsky, M.J.: Marketing and sustainability, Association with The Centre for Sustainability Design. Centre for Business Relationships, Accountability, Sustainability and Society (BRASS), UK (2002)
4. Chesbrough, H.: Managing Open Innovation. Research Technology Management 47(1), 23–26 (2004)
5. Dao, V., Langella, I., Carbo, J.: From green to sustainability: Information Technology and an integrated sustainability framework. Journal of Strategic Information Systems 20(1), 63–79 (2011)
6. Dewey, J.: How We Think. Dover Publication, Mineola (1997)
7. DiSalvo, C., Sengers, P.: Brynjarsdóttir. H.: Mapping the landscape of sustainable HCI. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2010), pp. 1975–1984. ACM, New York (2010)
8. DiSalvo, C., Sengers, P., Brynjarsdóttir, H.: Navigating the terrain of sustainable HCI. Interactions 17(4), 22–25 (2010)
9. Dourish, P.: HCI and environmental sustainability: the politics of design and the design of politics. In: Proceedings of the 8th ACM Conference on Designing Interactive Systems (DIS 2010), pp. 1–10. ACM, New York (2010)
10. Glavic, P., Lukman, R.: Review of sustainability terms and their definitions. Journal of Cleaner Production 15(18), 1875–1885 (2007)
11. Goodman, E.: Three environmental discourses in human-computer interaction. In: CHI EA 2009 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2009), pp. 2535–2544. ACM, New York (2009)

12. Huang, E.M., Truong, K.N.: Breaking the disposable technology paradigm: Opportunities for sustainable interaction design for mobile phones. In: Proceedings of the SIGCHI Conference on Human Factors in Computer Systems (CHI 2008), pp. 323–332. ACM, New York (2008)
13. Huh, J., Nam, K., Sharma, N.: Finding the lost treasure: Understanding reuse of used computing devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2010), pp. 1875–1878. ACM, New York (2010)
14. Kotler, P.: Marketing Management, 14th edn. Prentice-Hall (2011)
15. Mustaquim, M., Nyström, T.: Designing Sustainable IT System – From the Perspective of Universal Design Principles. In: Stephanidis, C., Antona, M. (eds.) UAHCI 2013, Part I. LNCS, vol. 8009, pp. 77–86. Springer, Heidelberg (2013)
16. Mustaquim, M.M., Nyström, T.: Design principles of open innovation concept: universal design viewpoint. In: Stephanidis, C., Antona, M. (eds.) UAHCI 2013, Part I. LNCS, vol. 8009, pp. 214–223. Springer, Heidelberg (2013)
17. Peattie, K., Peattie, S.: Social Marketing: a Pathway to Consumption Reduction? Journal of Business Research 62(2), 260–268 (2009)
18. Pierce, J., Brynjarsdottir, H., Sengers, P., Strengers, Y.: Everyday practice and sustainable HCI: understanding and learning from cultures of (un)sustainability. In: CHI 2011 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2011), pp. 9–12. ACM, New York (2011)
19. Porter, M.E., van der Linde, C.: Green and competitive: Ending the Stalemate. Harvard Business Review 73(5), 120–134 (1995)
20. Remy, C.: Taking a note from marketing research in sustainable HCI. Interactions 20(3), 12–15 (2013)
21. Schaltegger, S., Wagner, M.: Sustainable entrepreneurship and sustainability innovation: categories and interactions. Business Strategy and the Environment 20(4), 222–237 (2011)
22. Trott, P., Hartmann, D.: Why ‘open innovation’ is old wine in new bottles. International Journal of Innovation Management 13(4), 715–736 (2009)
23. United Nations: Report of the World Commission on Environment and Development, General Assembly Resolution 42/187 (1987)

Proactive Recommendation System for m-Tourism Application

Alexander Smirnov[1,2], Alexey Kashevnik[1], Andrew Ponomarev[1], Nikolay Shilov[1], and Nikolay Teslya[1]

[1] St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 39, 14 Line, 199178 St.Petersburg, Russia
[2] ITMO University, 49, Kronverkskiy pr., 197101 St. Petersburg, Russia
{smir,alexey,ponomarev,nick}@iias.spb.su

Abstract. In m-tourism applications, the proactive recommendations are especially actual for two major reasons: (1) the highly dynamic nature of the problem situation (the user continuously moves, the transport situation and weather conditions change); (2) limited possibilities of mobile devices for explicit information entry and checking large amounts of alternative solutions, but rich possibilities for tacit information entry via various sensors. The paper proposes an approach and research prototype based on the technologies of smart space and proactive recommendation systems. The architecture is based on the smart space technology. The system implementing the proposed approach helps the tourists to plan their attraction attending schedule based on the context information about the current situation in the region, its foreseen development, the tourist's preferences and previous behavior, using their mobile devices.

Keywords: m-tourism, infomobility, proactive recommendation system, smart space.

1 Introduction

The ubiquitous world in which we live is characterized by a high mobility of individuals, most of them wearing devices capable of geo-localization (smartphones or GPS-equipped cars) [1]. Mobiquitous environment is a next generation of ubiquitous environment, which supports adaptation to mobility of people and applications, and changes in devices state. In other words, mobiquitous environment has a mobile and ubiquitous nature. One of well-known technology that implements mobiquitous environment concept is the smart space. The smart space is an aggregation of devices, which can share their resources (information and services) and operate in coalitions. This nature of smart space enables of appearance of cyber conflicts between different smart space devices (or participants) which can have different goals and situation understanding but common information space for trusted cyber relationships.

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 113–127, 2014.

This is also a step to the "infomobility" infrastructure, i.e. towards operation and service provision schemes whereby the use and distribution of dynamic and selected multi-modal information to the users, both pre-trip and, more importantly, on-trip, play a fundamental role in attaining higher traffic and transport efficiency as well as higher quality levels in travel experience by the users [2]. It is a new way of service organization appeared together with the development of personal mobile and wearable devices capable to present user multimodal information at any time. Infomobility plays an important role in the development of efficient transportation systems, as well as in the improvement of the user support quality. In accordance with the forecast of [3], the market of such technologies as mobile Internet, automation of knowledge work, and Internet of Things by 2025 can increase 20 trillion USA dollars.

Another quickly developing concept is recommendation systems. This is a class of decision support systems that narrow down and rank the list of possible alternative solutions via estimating their utility. The information used for the expected utility estimation defines the class of the recommendation system and can include: user characteristics, history of previously made decisions, characteristics of the alternative solutions, current situation parameters, decision-making rules, etc.

Modern recommendation systems mostly work on the scheme “request-response”; the active party in the interaction is the user, who makes a request, which can contain some additional constraints for the alternative solutions. As a response, the recommendation system offers a list of recommendations (recommended solutions). In mobile recommendation systems, users cannot browse through many search results and suffer from other restrictions in the user experience, because of limitations in the user interface such as small display sizes or missing keyboards. In mobile environments, having the user not to submit any request or query to get a recommendation could possibly improve the user experience [4].

At the moment, there are hardly any systems that could offer recommendations in the proactive way, without user request, on the basis of the current situation and user profile analysis [5]. One of the reasons of this situation is the high risk of obtrusive offering non-relevant information to the user. However, modern mobile devices equipped with various sensors, make it possible to produce proactive recommendations, which would be useful and convenient due to creation of more precise user behavior models.

In m-tourism applications, the proactive recommendations are especially actual for two major reasons: (1) due to the highly dynamic nature of the problem situation (the user continuously moves, the transport situation and weather conditions change); (2) mobile devices usually have limited possibilities of explicit information entry and checking a large amount of alternative solutions, but rich possibilities for tacit information entry via various sensors.

Development of tourist services and applications is popular at the moment. “In a field trial in Görlitz (Germany), 421 tourists explored the city with one of two different mobile information systems, a proactive recommender of personalized tours and a pull service presenting context-based information on demand. A third group of tourists was tracked by GPS receivers during their exploration of the destination relying on traditional means of information. Results point out that both mobile applications gained a high level of acceptance by providing an experience very similar to a traditional guided tour. Compared to the group tracked by GPS loggers, tourists using a mobile information system discovered four times more sights and stayed at them twice as long." [6]

"The findings of the evaluation carried out have demonstrated that the widget-based solution is better than the notification-based solution. Despite the fact that both options are considered good solutions to achieve proactivity, the second one is considered by the users more annoying. <...> We can state that the "time pressure" factor is a good indicator to know when a proactive recommendation is reasonable or not, because in these situations users give less feedback." [7]

However, analysis of existing at the moment applications in the market shows that there is no application that can provide proactive tourist support based on his/her location, preferences, and current situation in the area (weather, traffic jams, and etc.). Hence, it is an actual task to develop an application that provides the tourist with the following information using personal mobile device:

- generate recommended attractions and their visiting schedule based on the tourist and region contexts and attraction estimations of other tourists; tourist context characterizes the situation of the tourist, it includes his/her location, co-travelers, and preferences; region context characterizes the current situation of tourist location area, it includes his/her location, co-travelers, and preferences; region context includes such information as weather, traffic jams, closed attraction, etc.
- collect information about attractions from different sources and recommend the tourist the best for him/her attraction images and descriptions;
- propose different transportation means for reaching the attraction;
- update the attraction visiting schedule based on the development of the current situation.

The paper proposes an approach and research prototype based on the technologies of smart space and proactive recommendation systems and implementing the above requirements. The next section describes the state-of-the-art in the area of e-tourism applications and recommendation systems. Section 3 describes the architecture of the presented application. The case study and developed prototype are presented in section 4. Major results are summarized in the conclusion.

2 Related Work

2.1 m-Tourism Applications

There is a number of projects aimed at tourist support depending on the current situation.

Smart Travelling [8] is an online travel guide that supports about 30 cities worldwide including the most interesting capitals of European countries and USA. The guide includes a database of restaurants, cafes, hotels, shopping-tips and other places of interests. The mobile application for iPhone is accessible through AppStore.

Integration with Google maps allows user to see the current location in the map and helps to navigate to each and every tip in destination cities. Application allows the user to download the content and use guide without Internet connection.

The travel guide Triposo [9] is a free mobile guide service available for Apple and Android devices. A user can download the application and appropriate database (which is updated ones each two months) to the mobile device beforehand and use it during the trip without Internet connection. The application supports logging of travelling. It includes databases from the following sources World66, Wikitravel, Wikipedia, Open Street Maps, TouristEye, Dmoz, Chefmoz and Flickr [10]. Each guide contains information on sightseeing, nightlife, restaurants and more.

GoTour [11] is an Android-based mobile application for providing tourism and geographic services in Istanbul city. Application has internal attraction database and provides possibilities of searching places of interests around using the Variable Neighborhood-based algorithm.

Millions of traveler reviews, photos, and maps can be accessible in TripAdvisor [12]. Tourists can plan their trips taking into account over 100 million reviews and opinions by travelers. TripAdvisor makes it easy to find the lowest airfare, best hotels, great restaurants, and fun things to do, wherever you go. The mobile application is free, it supports all mobile platforms.

The Tourist Attractions [13] system determines the location and implements search for places of interests nearby in Google.

The Foursquare application [14] shows places of interests around the tourist in the interactive map and provides possibilities to estimate attended places. In addition, application allows to check-in at the attraction for showing the tourist location to other application users.

Smart Museum application [15] provides possibilities to see information about places of interests from a prepared beforehand database. Application also provides possibilities of getting information inside museums using the Internet connection and special Smart Box installed by the Smart Museum company in different museums.

The carried out analysis of the mentioned above systems shows that they can be divided into three main groups:

- applications that implement search for information around the tourist (e.g., World Explorer uses Wikipedia database, Tourist attractions uses Google database);
- applications that have own databases with information about attractions and provide this information to the user (e.g., SmartMuseum, Triposo);
- applications that collect estimation information about attraction estimations and suggest tourists if this attraction is good or not (e.g., Foursquare, Tripadvisor).

None of the existing applications gives recommendations based on the foreseeing of the current situation development and user behavior, i.e., they are not proactive.

2.2 Smart Spaces

Since proactive recommendation systems for m-Tourism applications extensively use various information sources and computational resources, it is reasonable to implement them based on the service-oriented architecture, and the integration of mobile devices can be supported by the concept of smart spaces. Smart spaces as a one of the major technologies of the Internet of Things is an integration of services provided by various devices so that they could share information for implementation of the ubiquitous computing paradigm. The concept of smart spaces makes it possible to re-distribute tasks depending on the types of devices solving them, and computational capabilities of these devices.

In modern projects, related to smart spaces, the main attention is paid to building intelligent systems, i.e. systems aimed at knowledge processing. These systems use such technologies as ontology and context management, profiling, knowledge processing and representation, etc.

For example, in [16] a concept of Smart-M3 information sharing platform is presented. This platform enables information and knowledge sharing between mobile devices in a smart space. It is based on the smart space RDF ontology and SSAP (smart space access protocol). In [17] a multi-agent architecture is presented, which integrates ontological model of context description and logic-based model into a first-order probabilistic logic-based model. This made it possible to describe the basic structure of the context and design the mechanism of probabilistic reasoning. A multi-agent architecture and smart space ontology COSE (Casas Ontology for Smart Environments) is presented in [18].

2.3 Proactive Recommendation Systems

Recommendation systems are widely used in the Internet for suggesting products, activities (including tourism), etc. for a single user considering his/her interests and tastes (e.g., [19]). The collaborative / group recommendation systems try to find users who share similar interests with the given user and recommend items they choose to that user [20]. However, the collaborative recommendation is complicated by the necessity to take into account not only personal interests but to compromise between the group interests and interests of the individuals of this group.

Until recently, the most recommendation systems operated in the 2-dimentional space "user-product". They did not take into account the context information, which, in most applications can be critical. As a result, there was a need in development of group recommendation systems based not only on previously made decisions but also on the contexts of situations in which the decisions were made. This gave a rise to development of context-driven collaborative algorithms of recommendation generation since their usage would significantly increase the quality and speed of decision making.

The recommendation systems can be split into the following classes based on the level of user involvement into the process of recommendation delivery:

- Pull-based. Reply to the user's queries and requests (less obtrusive).
- Reactive. React to changing situational context to generate content recommendations.
- Proactive. Proactively precaching appropriate content through extrapolating future context using specialized prediction models.

Since proactive recommendation systems form a new research field, in this subsection a state-of-the-art in the areas, methods and models from which can be used for development of such systems, is presented.

The proactivity mechanisms in various information systems have been discussed for years already. For example, the system XLibris [21] is aimed at searching additional possibly useful for the user documents, while he/she is working on a document.

Appearance of mobile devices caused a more intensive development of proactive systems. The most popular idea is presenting information, when the user is located in a certain place or near it. For example, CyberMinder [22] displays previously set by the user reminders associated with a certain place. Usage of this idea facilitated the appearance of various automated tourist guides.

The next step in the development of proactive systems is appearance of systems based on prediction of the user behavior, his/her future locations [23] or actions [24, 25], as well as situation recognition [26]. Today, integration of proactive systems with recommendation systems can be considered as a perspective research field [27; 28].

3 m-Tourism Proactive Recommendation System

3.1 Approach

The proposed approach to building proactive recommendation system is based on the following principles:

- **Application of ontologies for problem domain description.** Usage of ontologies [29] enables complementing user descriptions with relations to ontology concepts. This, in turn, will make it possible to improve recommendations based on usage of the relationships between ontology concepts.
- **Context-dependence.** Current situation modelling with the help of context management technology enables selecting only information, which is needed at the given moment of time, and hence, increases the quality of generated recommendations
- **Usage of social media for user profiling.** The information produced during the interaction of the user with social networks can be treated as semi-structured personal information space, which characterizes the user's interests and preferences.
- **Proactive system behavior.** The analysis of the decisions made by the user and other users with similar characteristics (interests, preferences, etc.) earlier and their explicit and tacit preferences makes it possible to foresee their needs taking into account possible development of the current situation and to produce recommendations not by user request but proactively shifting the level of intelligent decision support quality.

The user's behavior prediction is based on the elements of the theory of planned behavior [30]. In accordance with this theory, the behavior is mostly defined by the following factors, which have to be taken into account in proactive intelligent decision support:

- **Action.** Defines the type of behavior (e.g., social behavior, political behavior, social interaction, etc.).
- **Object.** Defines the behavior object (social service, consumer product, relative, etc.).
- **Context.** Defines the context of the behavior (social-economic situation, in conditions of insufficient resources or sufficient resources, on public or in an intimate setting, etc.).
- **Time factor.** The time of action is analyzed: e.g., right away, in a year, during several years, at a certain date, etc.

The functional behavior analysis is carried out after revealing the key behavior related to the human activity under consideration. The functional behavior analysis is often referred to as ABC-analysis (Antecedent, Behavior, Consequence) and is based on revealing of both antecedents and consequences of the particular behavior.

As a result, it is possible to build a conditional human (user) behavior model, which would make it possible to analyze how human behavior appears, carried out and regulated to better predict his/her actions and decisions.

3.2 Architecture and Scenario

The proposed intelligent mobile tourist guide consists of the following main components (Fig. 1):

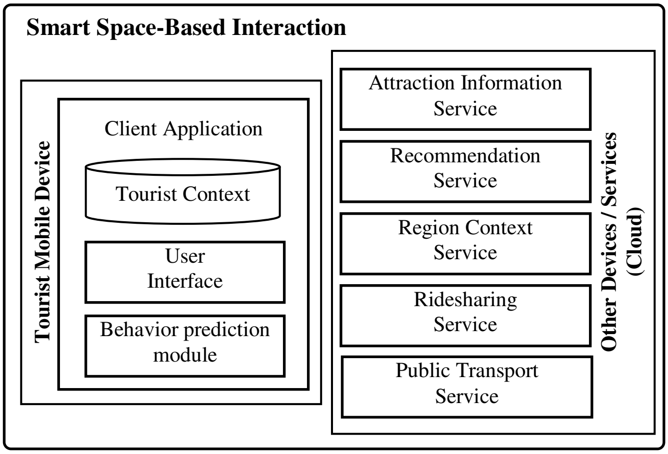


Fig. 1. General architecture of intelligent mobile tourist guide

- client application installed on the user mobile device that shares tourist context with the smart space, tries to foresee (predict) user actions for generating pro-active recommendations, and provides the tourist results of guide application operation;
- attraction information service that retrieves and caches information about attractions;
- recommendation service that evaluates attraction scores based on ratings saved earlier [31];
- region context service that acquires and provides information about the current situation in the considered region (e.g., weather, traffic jams, closed attractions);
- ridesharing service that finds matching routes of the tourist and preferred attraction locations with accessible in the region drivers routes; the service provides the tourist possibilities of comfort transportation to preferred attractions [32];
- public transport service that finds information about public transport applicable to reach the preferred attraction [33].

The client application is located on the user mobile device while other services use powerful computer systems. As a result, the main engineering task of the mobile tourist guide development is implementation of resource-intensive operations using other mobile tourist guide services, which is located in computational servers. Thereby, the main tasks of client application are: share and process information about tourist context, profile, and actions; communicate with smart space; provide results to the tourist; share tourist ratings of attended attractions; browse descriptions and images of the attractions.

The attraction information service is responsible for six main tasks. The first one is extraction attractions titles around the tourist from external sources and shares it in the smart space. The second task is caching acquired information for quick access to external sources by tourists in the same location without additional requests. The third task is providing internal identifiers for attractions, which are stored in Geo2Tag platform [34] (the platform provides possibilities for quick search of identifiers by locations). The fourth task is extraction of default images for attractions that are stored in the internal database. Default images are defined by the recommendation service as the best images for an attraction based on tourists ratings. The fifth task is extraction of attraction details (lists of images and text descriptions) from external sources and sharing them with the smart space. The sixth task is to set and refresh default images in the internal database for the attractions based on the information available in the smart space from the recommendation service.

The recommendation service is responsible for ranking attractions, images, and descriptions for providing the tourist the best attractions to see and the best images and description of chosen attraction for acquaintance. It stores the following mappings in the internal database: (internal attraction identifier, tourist context, region context -> attraction rating), (image URL -> image rating), (description URL, description text -> description rating). The internal database also keeps similarities between users, which are calculated as background process.

The region context service acquires information about tourist location area from region specific services as soon as the client application publishes the tourist context to the smart space. The context information is published to the smart space and used by other services.

The public transport and ridesharing services provide alternatives for transportation means for the tourist to reach preferred attractions. While the public transport service provides the tourist a route that describes the sequence of transportation means to reach the preferred attraction, the ridesharing service finds a driver, which goes from the tourist location to the same direction.

4 Prototype and Case Study

Implementation of the intelligent mobile tourist guide is based on the Android OS that is one of the most popular operating systems for mobile devices today [35]. The current version of the guide (which does not include proactive behavior) is available for free downloading in Google Play Market[1]

Fig. 2 shows the main application screen and settings page. In the main screen the tourist can see images extracted from accessible internet sources around, clickable map with his/her location, context situation, and the best attractions around ranked by the recommendation service. By pressing “menu” button guide application allows to search information for worldwide attractions by choosing another area (country, region, and city) and access the settings page of the mobile tourist guide application. In the status bar the tourist can search for attractions worldwide.

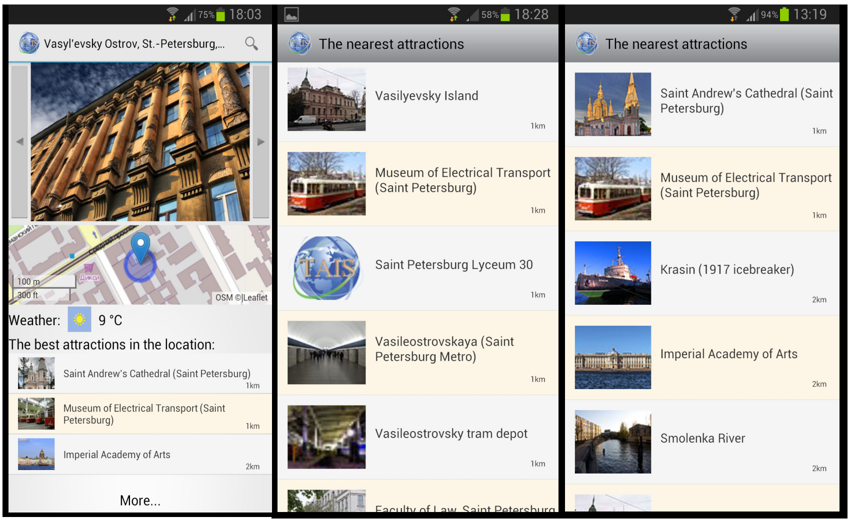


Fig. 2. Main application screen (left), list of attractions with recommendations (middle) and without (right)

---

1 Mobile Tourist Guide Client for Android OS is available at `https://play.google.com/store/apps/details?id=ru.nw.spiiras.tais`

In the settings screen the following parameters can be chosen: language, recommendations using, and smart space connection settings (for using recommendation and other services). Parameter language determines the language of the extracted information, while the interface language is determined by the Android OS settings.

Tourist can browse information about the best attractions around presented by the mobile tourist guide in the main screen and press button “More” to see more attractions (see Fig. 2, left).

In the left screenshot a list of attractions without using recommendation service is shown (attractions are ordered by distance from the user). In the right screenshot the attractions are ordered by the recommendation service according to attraction estimations. Table 1 accumulates information about previous estimations of the tourist (John) and other tourists (Lisa, Alice, Bob) who have similar interests.

As it can be seen from Table I the mobile tourist guide application users (Lisa, Alice, Bob, Patrick, Jack) have estimations of attractions around the tourist (Museum of Electrical Transport, Imperial Academy of Arts, Krasin, Smolenka River, Saint Andrew’s Cathedral). The tourist (John) doesn’t have estimations for these attractions because he has never been there before. John, Lisa, Alice, and Bob have been integrated into a group based on the information about their previous estimations of the same attractions.

Then, the recommendation service determine estimations of all users in this group for attractions nearby the tourist (see Table 2). Based on these estimations, similarities between tourists, distances to attractions, and context information, the recommendation service sorts attractions around the tourist (see Fig. 2 right).

Table 1. Example of attraction estimations of mobile tourist guide application users

	Louvre	Eiffel Tower	Palace of Versailles	Notre Dame de Paris	Ile de la Cite	Museum of Electrical Transport	Imperial Academy of Arts	Krasin	Smolenka River	Saint Andrew’s Cathedral
John	*5*	*5*	*3*	*3*	*2*					
Lisa		5	3	2	2	**5**	**5**		**5**	**5**
Alice	5	5	3	2	2	**4**	**5**	**5**	**4**	**5**
Bob	5	5	3	2	2	**5**	**4**	**5**		**5**
Patrick	2	1		3	4	5	2	4		2
Jack		4	5		5		3	4		3

Table 2. Group of guide application users with similar estimations as the tourist

	Museum of Electrical Transport	Imperial Academy of Arts	Krasin	Smolenka River	Saint Andrew’s Cathedral
Lisa	5	5		5	5
Alice	4	5	5	4	5
Bob	5	4	5		5

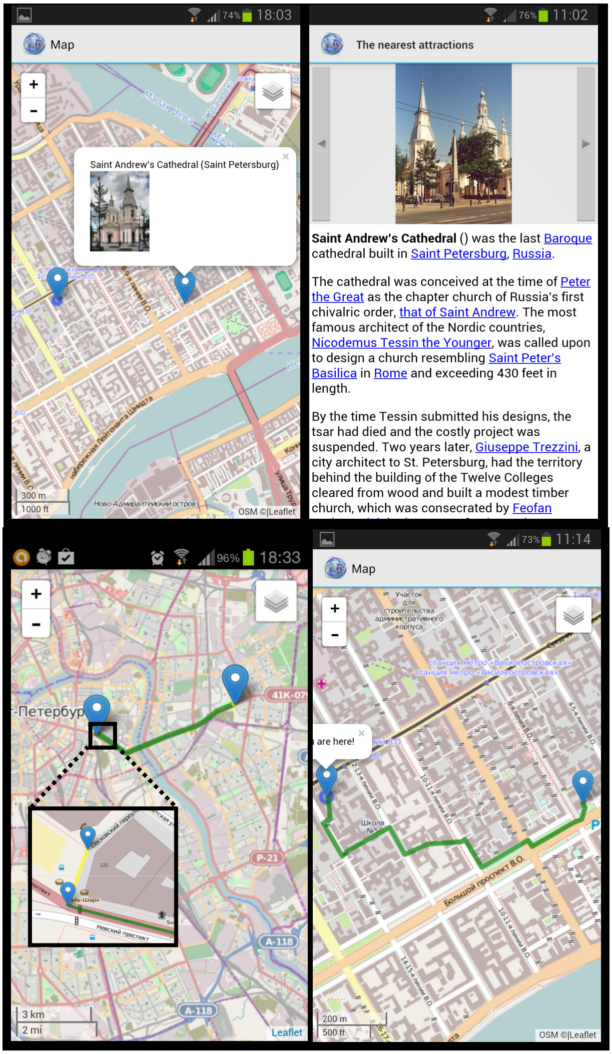


Fig. 3. An attraction in the interactive map (top left), attraction description page (top right), finding drivers for tourists that go to the same direction (bottom left and right)

The tourist has possibilities to see recommended attractions in the interactive map (Fig. 3, top left screenshot). OpenStreetMap is used as a base platform for maps operations [36]. In Fig. 3 (the top right screenshot) images and description for attraction are shown. The tourist can leaf over images to browse them. Images are extracted from different Internet sources ranked by the recommendation service according to their estimations.

The behavior prediction module accumulates such information as usual user's schedule (time when the user wakes up, eats, goes to bed), most visited brands of restaurants, as well as preferred ways of transportation (bus, train, car, walking). Based on this information, the behavior prediction module analyses the solutions

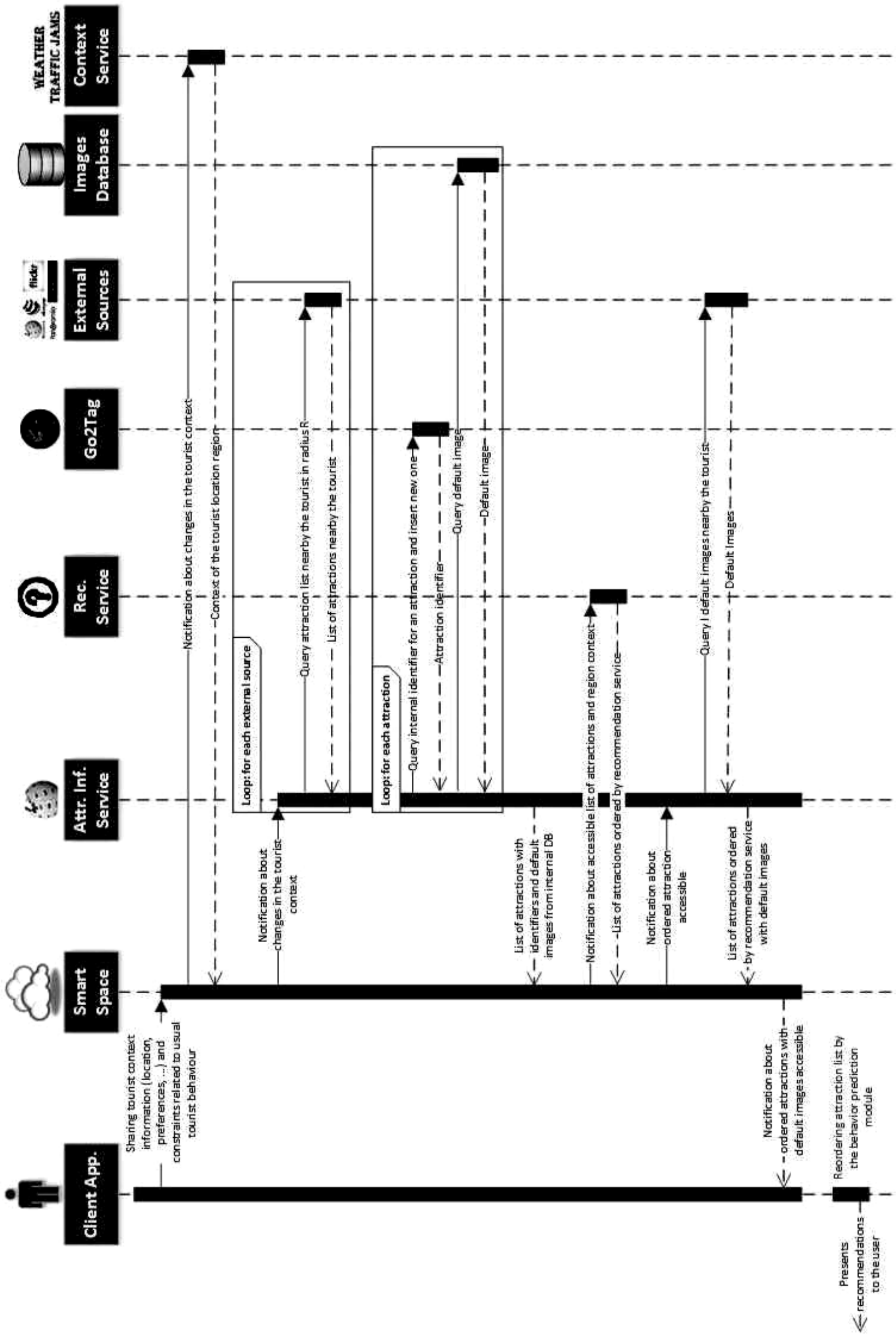


Fig. 4. Service interaction scenario

comparing them with revealed user preferences and modifies the list of the preferred attractions or provides corresponding proactive recommendations. For example, in the situation, when the user is planning to visit a museum, which is reachable in 1 hour and its usual attendance is estimated as three hours, and the user is used to have a lunch in two hours; the application would recommend another attraction instead, so that the user could have a lunch at usual time in a preferred brand of restaurants.

After the final choice is made by the user, the ridesharing service finds a driver for the tourist to reach the attraction. The tourist interactively gets information about possible driver with his/her profiles, meeting points, meeting time, full recommendations about the route (Fig. 3, bottom left screenshot). The path in the figure means that the tourist has to walk to the meeting point with the driver, then he/she goes with the driver, leaves the car at the drop-off point, and then walks to the attraction. Bottom right screenshot in Fig. 3 shows pedestrian path from the tourist location to the interesting attraction.

The overall scenario of the above described service interaction is shown in Fig. 4.

The developed context service searches for weather conditions in a user location at the moment. It uses World Weather Online [37] resource to get this information. The resource API allows to implement not more than 500 requests per hour for free and has premium API for an increased number of requests. Weather information is updated every 3-4 hours.

5 Conclusions

The paper proposes an architecture and a research prototype of proactive recommending system for m-Tourism. The system helps the tourists to plan their attraction attending schedule based on the context information about the current situation in the region, its foreseen development, the tourist's preferences and previous behavior, using their mobile devices.

The project is still under development. The prototype available doesn't support proactive recommendations yet. Besides, if the user wishes to adjust generated solutions via adding additional constraints or preferences the solutions will have to be generated from scratch (only the context is reused). Future work will address incremental solution adaptation via iterative interaction with the user based on planning models (e.g., opportunistic planning [38]) along with the proactive recommendations.

Acknowledgments. The research was supported partly by projects funded by grants 13-07-00271, 13-07-00039, 13-07-12095, 14-07-00378, and 13-07-00336 of the Russian Foundation for Basic Research, project 213 (program 15) of the Presidium of the Russian Academy of Sciences, and project #2.2 of the basic research program "Intelligent information technologies, system analysis and automation" of the Nanotechnology and Information technology Department of the Russian Academy of Sciences. This work was also partially financially supported by Government of Russian Federation, Grant 074-U01.

References

1. Artigues, C., Deswarte, Y., Guiochet, J., et al.: AMORES: an architecture for mobiquitous resilient systems. In: ARMOR 2012 Proceedings of the 1st European Workshop on AppRoaches to MObiquiTous Resilience, pp. 7:1–7:6. ACM, New York (2012)

2. Ambrosino, G., Nelson, J.D., Bastogi, B., Viti, A., Romazzotti, D., Ercoli, E.: The role and perspectives of the large-scale Flexible Transport Agency in the management of public transport in urban areas. In: Ambrosino, G., Boero, M., Nelson, J.D., Romanazzo, M. (eds.) Infomobility Systems and Sustainable Transport Services, pp. 156–165. ENEA, Rome (2010)
3. McKinsey Global Institute, `http://www.mckinsey.com/`
4. Woerdnl, W., Huebner, J., Bader, R., Gallego-Vico, D.: A Model for Proactivity in Mobile, Context-Aware Recommender Systems. In: Proceedings of the Fifth ACM Conference on Recommender Systems, pp. 273–276 (2011)
5. Ricci, F.: Mobile Recommender Systems. J. of IT & Tourism 12, 205–231 (2011)
6. Modsching, M., Kramer, R., ten Hagen, K., Gretzel, U.: Effectiveness of mobile recommender systems for tourist destinations: A user evaluation. In: Yorke-Smith, N. (ed.) Interaction Challenges for Intelligent Assistants: Papers from the AAAI Spring Symposium. Technical Report SS-07-04, pp. 88–89 (2007)
7. Bader, R., Siegmund, O., Woerndl, W.: A Study on User Acceptance of Proactive In-Vehicle Recommender Systems. In: 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2011), Salzburg, Austria, pp. 47–54 (2011)
8. Smart Travelling web page, `http://www.smart-travelling.net/en/`
9. "New travel guide Triposo brings algorithms to apps," The Independent (September 2011), `http://www.independent.co.uk/travel/news-and-advice/new-travel-guide-triposo-brings-algorithms-to-apps-2353352.html`
10. Triposo, `http://www.triposo.com`
11. Al-Rayes, K., Sevkli, A., Al-Moaiqel, H., Al-Ajlan, H., Al-Salem, K., Al- Fantoukh, N.: A Mobile Tourist Guide for Trip Planning. IEEE Multidisciplinary Engineering Education Magazine 6(4), 1–6 (2011)
12. Tripadvisor, `http://www.tripadvisor.ru`
13. Google Play Market, Tourist Attractions application, `https://play.google.com/store/apps/details?id=com.uknowapps.android.attractions&hl=ru`
14. Foursquare application, `http://foursquare.com`
15. Smart Museum, `http://smartmuseum.ru/`
16. Honkola, J., Laine, H., Brown, R., Tyrkko, O.: Smart-M3 Information Sharing Platform. In: Proc. IEEE Symp. Computers and Communications (ISCC 2010), pp. 1041–1046. IEEE Comp. Soc. (2010)
17. Weijun, Q.I.N., Yuanchun, S.H.I., Yue, S.U.O.: Ontology-based context-aware middleware for smart space. Journal of Tsinghua Science and Technology 12(6), 707–713 (2007)
18. Wemlinger, Z., Holder, L.: The COSE Ontology: Bringing the Semantic Web to Smart Environments. In: Abdulrazak, B., Giroux, S., Bouchard, B., Pigot, H., Mokhtari, M. (eds.) ICOST 2011. LNCS, vol. 6719, pp. 205–209. Springer, Heidelberg (2011)
19. Garcia, I., Sebastia, L., Onaindia, E., Guzman, C.: A Group Recommender System for Tourist Activities. In: Di Noia, T., Buccafurri, F. (eds.) EC-Web 2009. LNCS, vol. 5692, pp. 26–37. Springer, Heidelberg (2009)
20. Kumar, A., Thambidurai, P.: Collaborative Web Recommendation Systems Based on an Effective Fuzzy Association Rule Mining Algorithm (FARM). Indian Journal of Computer Science and Engineering 1(3), 184–191 (2010)
21. Price, M., Golovchinsky, G., Schilit, B.N.: Linking By Inking: Trailblazing in a Paper-like Hypertext. In: Conference on Hypertext and Hypermedia, Pittsburgh, PA, USA, pp. 30–39 (1998)
22. Dey, A.K., Abowd, G.D.: CyberMinder: a context-aware system for supporting reminders. In: Symposium on Handheld and Ubiquitous Computing, Bristol, UK, pp. 172–186 (2000)

23. Hinze, A., Sachs, K., Buchmann, A.: Event-based applications and enabling technologies. In: Proceedings of the Third ACM International Conference on Distributed Event-Based Systems, Nashville, TN, USA, pp. 1:1–1:15 (2009)
24. Bellotti, V., Price, B., Rasmussen, P., Roberts, M., Schiano, D.J., Walendowski, A., Begole, B., Chi, E.H., Ducheneaut, N., Fang, J., Isaacs, E., King, T., Newman, M.W., Partridge, K.: Activity-based serendipitous recommendations with the Magitti mobile leisure guide. In: Conference on Human Factors in Computing Systems, New York, NY, USA, pp. 1157–1166 (2008)
25. Partridge, K., Price, B.: Enhancing Mobile Recommender Systems with Activity Inference. In: Houben, G.-J., McCalla, G., Pianesi, F., Zancanaro, M. (eds.) UMAP 2009. LNCS, vol. 5535, pp. 307–318. Springer, Heidelberg (2009)
26. Ciaramella, A., Cimino, M.G.C.A., Marcelloni, F., Straccia, U.: Combining Fuzzy Logic and Semantic Web to Enable Situation-Awareness in Service Recommendation. In: Bringas, P.G., Hameurlain, A., Quirchmayr, G. (eds.) DEXA 2010, Part I. LNCS, vol. 6261, pp. 31–45. Springer, Heidelberg (2010)
27. Worndl, W., Brocco, M., Eigner, R.: A Context-Aware Gas Station Recommender System for Vehicular Ad-Hoc Networks. In: Conference on Wireless Applications and Computing Conference, Amsterdam, Netherlands, pp. 101–108 (2008)
28. Turlier, S., Hahn, C., Gebhardt, S.: Browsing online music catalogs in a vehicle. In: Workshop on Mobile Cloud Media Computing, Firenze, Italy, pp. 53–58 (2010)
29. Smirnov, A., Levashova, T., Shilov, N.: Semantic-oriented support of interoperability between production information systems. International Journal of Product Development 4(3/4), 225–240 (2007)
30. Fishbein, M., Ajzen, I.: Belief, attitude, intention, and behavior: An introduction to theory and research. Addison-Wesley, Reading (1975)
31. Smirnov, A., Kashevnik, A., Ponomarev, A., Shilov, N., Shchekotov, M., Teslya, N.: Recommendation System for Tourist Attraction Information Service. In: Proc. FRUCT Conf., pp. 140–147 (November 2013)
32. Smirnov, A., Shilov, N., Kashevnik, A., Teslya, N., Laizane, S.: Smart Space-based Ridesharing Service in e-Tourism Application for Karelia Region Accessibility. Ontology-based Approach and Implementation. In: Proc. 8th Int. Joint Conference on Software Technologies, Reykjavik, Iceland, July 29-31, pp. 591–598 (2013)
33. Smirnov, A., Kashevnik, A., Balandin, S.I., Laizane, S.: Intelligent Mobile Tourist Guide: Context-Based Approach and Implementation. In: Balandin, S., Andreev, S., Koucheryavy, Y. (eds.) NEW2AN/ruSMART 2013. LNCS, vol. 8121, pp. 94–106. Springer, Heidelberg (2013)
34. Krinkin, K., Yudenok, K.: Geo-coding in Smart Environment: Integration Principles of Smart-M3 and Geo2Tag. In: Balandin, S., Andreev, S., Koucheryavy, Y. (eds.) NEW2AN/ruSMART 2013. LNCS, vol. 8121, pp. 107–116. Springer, Heidelberg (2013)
35. International Data Corporation (IDC), Press Release, Android and iOS Combine for 91.1% of the Worldwide Smartphone OS Market in 4Q12 and 87.6% for the Year, `http://www.idc.com/getdoc.jsp?containerId=prUS23946013` (last access date: April 30, 2014)
36. Smirnov, A., Shilov, N., Kashevnik, A., Teslya, N.: OpenStreetMap-Based Dynamic Ridesharing Service. In: Proc. Information Fusion and Geographic Information Systems, Sixth International Workshop, pp. 103–118 (May 2013)
37. World Weather Online, `http://worldweatheronline.com`
38. Hayes-Roth, B.: Human Planning Processes. Scientific Report (1980), `http://www.rand.org/content/dam/rand/pubs/reports/2007/R2670.pdf`

What Capability Is Not

Anders W. Tell

Stockholm University,
Department of Computer and Systems Sciences, Sweden
anderswt@dsv.su.se

Abstract. There is no broad agreement of the nature of capability in management, planning, engineering, and enterprise architecture literature. Definitions of the concept of capability ranges from being equal to the concept of process to be loosely defined as a collection of resources. This paper presents intensional aspects that are argued to be non-essential for the general concept of capability. Thus, this paper explores the concept of capability with a focus on what capability is not. By making non-essential aspects and distinctions visible we aim to making the concept of capability easier to understand and integrate for practitioners and researchers.

Keywords: Capability, ability, capability management, capability planning, resource, function, process.

1 Introduction

In this paper the concepts of capability and implicitly ability are explored. The words capability and ability has been part of natural language since 1300-1400 [1] [2] and the underlying concepts are being used to represent the inner workings of societies, enterprises, organisations, man-made artefacts, biological systems and organisms.

During the last decade, capability constructs are found beside, on top of, or as replacement of well-known constructs, theories and paradigms used to represent (natural, social, man made, biological) systems. Examples of overlapping constructs include process, component, goal, function, resource, resource based theory of the firm. As such the capability construct can be considered as the "The new kid in the (framework building) block", for solving a wide range of problems.

What can be noted is the emergence of a great number of conceptualisations of capability. Dosi, Nelson and Winter stated in their book: "The Nature and Dynamics of Organizational Capabilities" [3]:

> "The term 'capabilities' floats in the literature like an iceberg in a foggy Arctic sea, one iceberg among many, not easily recognized as different from several icebergs near by."

The proliferation of definitions and usages show few signs of slowing down. The following examples from different domains may aid the understanding of the variety of definitions and paradigms.

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 128–142, 2014.

Competence Based Management:

"repeatable patterns of action that are created through a firms management processes for coordinating its resource in processes for value creation." Ron Sanchez [4].

Military Domain:

"DoDAF: The ability to achieve a desired effect under specified [performance] standards and conditions through combinations of ways and means [activities and resources] to perform a set of activities" UPDM v2 [5].

Information Technology Strategy and Management:

"the ability to marshal resources to affect a predetermined outcome" McKeen and Smiths [6].

Service Oriented Architecture:

"Capability is a real-world effect that a service provider is able to provide to a service consumer", and their related Service concept" OASIS SOA Reference model [7] .

Human Development:

”The capability approach involves ’concentration on freedoms to achieve in general and the capabilities to function in particular’ (Sen 1995). The major constituents of the capability approach are functionings and capabilities. Functionings are the ’beings and doings’ of a person, whereas a person’s capability is ’the various combinations of functionings that a person can achieve. Capability is thus a set of vectors of functionings, reflecting the person’s freedom to lead one type of life or another’ (Sen 1992).” [8].[1]

ISO 9000:

“ability of an organization, system or process to realize a product that will fulfil the requirements for that product”

The on-going debate illustrates a proliferation of capability definitions, theories, and approaches with no broad agreement on the general nature. The research question addressed in this paper is whether the lack of broad consensus can be addressed by seeking agreement on "what capability is not". Can the identification and exploration of non-essential aspects and distinctions provide a baseline for "is not" and "is" discussions and agreements? This question is part of a design science inquiry into the requirements that influence the design of a knowledge organisation construct, Ability Perspective, i.e. an Architecture Viewpoint [9].

Knowing "what capability is not", and essential characteristics provides assistance, not only for definitional work but also for incorporation of the concepts into existing theories, frameworks and methodologies. Secondly, the process of comparing capability-oriented theories becomes easier when the underpinning assumptions are made visible, e.g. when a capability theory is assumed to be relevant (and valuable) only for organisations.

[1] This quote from Ingrid Robeyns include references to work of Nobel laureate Amartya Sen.

The aim of this paper is to identify and explore non-essential aspects and distinctions in order to further harmonisation-efforts amongst researchers and practitioners.

The approach of this paper is inspired by the work of Robert I. Sutton and Barry M. Shaw, in "What theory is not"[10], and Richard Baskerville in "What design science is not" [11], where they suggest that broader agreements may be easier achieved by seeking agreement on what a “theory” or “design science” is not. The informal argumentation is supported by the use of test capability cases (capability sentences), and a demonstration of one capability definition that is consistent with the explored non-essential aspects and distinctions.

The scope of this paper is a reduced (non-exhaustive) set of aspects and distinctions in order to keep the paper short. There is, of course, an infinite set of possible "is not" aspects.

The main contributions of this paper are the identified aspects, capability cases, and distinctions that are argued to be non-essential of the nature of common sense and general capability concepts.

The remainder of the paper is structured as follows: In section 2 the identified non-essential aspects, capability cases, and distinctions are presented together with argumentation. Section 3 briefly outlines candidate intensional parts of a general and common sense meaning of capability concepts. i.e. "what a capability is”. The paper is concluded with a conclusion section.

2 What Capability Is Not

In this section aspects and distinctions are presented that are argued to not characterise a general and common sense conceptualisation of capability. The following groups of aspects and distinctions are covered is this paper:

Table 1. Aspects and distinctions are covered is this paper

Group	Aspects and distinction
General distinctions	What-How, Abstract-Concrete
Ontological aspects	Context, Abstract-Concrete
Related concepts	Function, Resource, Process
Characteristics, modifiers, attitudes	Intention, Abstract
What can be capable	Organization
Work Perspectives	Kind, Context
Sources of capability	Process, Resource
Results of capability	Result

The list is by no means exhaustive; there exist an infinite set of possible "is not" aspects. For some of the presented aspects the subsection title is shortened with respect to the argumentation. For the "Capability is not a resource" section the argumentation suggests that not all capabilities are resources, although some are, i.e. Capability is not a proper subcategory of Resource.

2.1 Capability Is Not Characterised as What, as Opposed to How

The distinction between the interrogatives 'what' and 'how' is sometimes used to illustrate and define capability within the business architecture domain [12] and (software) engineering. Relying on the what-how distinction pose a number of practical problems since 'what' is polysynomous and 'how' is flexible.

In the field of Journalism the 5W interrogatives (what, who, why, when, or where) are recommended for coverage of a story. Sometimes +1H (how) is added as further guidance. On the other hand when used in software engineering the what-how often signify various other intuitions and distinctions such as requirement-realisation, external-internal, output/outcome-process, declarative-procedural knowledge, theoretical-practical knowledge, and everything else that are not details about the underlying process.

From a language point of view the interpretation and usage 'what' and 'how' often depend on a constructs including attitudes such as know ("know how-to"), belief, learn, ask, explain, etc. In an interesting turn of argumentation, ability is sometimes used as a basis for defining 'how-to' sentences, which suggests that ‘how’ can defined based on capability. Not all scholars agrees to this approach [13]. '

When considering a fuller international language perspective the introduction of lexical ambiguity leads to potential problems for non-native English speakers and practitioners.

From a pragmatic point of view answers to 'what' and 'how' questions vary across agents even in a single discourse context [13]. The same answer may be given to a ‘what’ and a ‘how’ question.

a) What happened? I mixed ingredients.

b) How did it happen? I mixed ingredients.

The possibility of interpreting the conceptual difference between what and how as an ontological difference between endurants/continuant and perdurant/occurrents, provides yet another dimension of separating what from how, leading to even more ambiguity relating to the distinction.

The seemingly simple and straightforward distinction between the interrogatives, what-how, does not provide unambiguous semantic definitional power and pragmatic aid for practitioners.

2.2 Capability Is Not an Organisational Construct

In the management field capability and competences are two concepts commonly attributed to organisations and corporations [14]. However the concept of capability is applicable in many other subject fields, biology, human resource, etc. In Financial Times one can find the following expressions:

a) Athens' ability to stay course in doubt

b) This EU move has the ability to take the City down

Here we find an example where a city (or possibly indirectly a country) may or may not be (cap-)able (adjective) 'to' something. In (c) a person is capable; in (d) a machine performer can change the economical world, and in (e) an organ can change a biological system.

c) Nils is capable to can bake healthy cakes that makes you happy.

d) The baking machine is capable to effectively utilize ingredients and produce marketable pastries.

e) The heart is capable to oxygenate blood.

A capability may be attributed to an organisation but other entities can possess/have capabilities. Therefore in general, a capability is not only an organisational construct. Within a specific subject field, theory or situation only organisational capabilities may be considered as relevant, although such specific restrictions should be motivated.

2.3 Capability Is Not of a Single Kind

In the same way that a process, or an asset can be viewed in many ways, such as concrete, high level, organisational, etc., there exist different categories of capability. It is easy to imaging that the nature of the possessor of a capability, and enclosing work perspective has an impact on which kinds of capabilities that are considered as relevant. From a management work perspective expressions (a, b) is relevant, (c) is relevant from a human resource work perspective, (d) is relevant from an economical theory point of view, and (e) relates to satisfying (abstract, experiential) customer value.

a) The capability of an organisation to sustain competitive advantage.

b) We plan to be capable to buy and integrate the capabilities that our business unit needs.

c) Nils is capable of applying psycho-motorical skills.

d) We can set the 'right' price

e) Capability to create a confortable home that customers desire to live in

A complicating factor when comparing capability approaches is that the "capability"-designation, or label, often is the same across approaches, but the definitions and their use are different. The practice of using the same label (term) for different kinds of capability is likely to prolong the process of harmonising approaches to capability.

There exists many kinds of capability that are interesting to knowledge workers and relevant to consider as part of a theory, framework or a methodology, and authors should be careful labelling capability constructs, declaring applicable work perspectives and situations.

2.4 Capability Is Not a Resource

During the later part of 2000th century the concept of resource obtains specific meanings in economical literature such as in the Resource Based View (RBV) of firms developed by Barney (1986, 1991, 1997) [4]. Here resource is defined as a quality of a firm:

> "(i) A resource is a ‘‘firm attribute’’ that is strategically valuable because it enables a firm to undertake actions in its product markets that improve the firm’s efficiency and effectiveness and thereby enable the firm to charge profitable prices for its products." [3, 4]

In many cases authors relate resources with capabilities [15], which pose an interesting question, are all capabilities also resources? In general, owning, having access to, or acquiring a resource opens up to the possibility of something of value to occur.

The case expressed in (a) involves a machine, a resource that is part of a value chain that leads-to sellable products. The experiential ability of the bakery in (b) leads to something valuable for the customer. However in (c) the same machine also give raise to something that is not valuable, not to the environment and not to the customer.

> a) Our machine is producing cakes that our customers love.
>
> b) The bakery is capable to enrich customer’s life.
>
> c) The bakery machine is capable to produce an abundance of grease residuals and exhausts that pollute the environment, which our customers do not like.

There exist capabilities that do not lead-to something valuable, i.e. do not pass RBV criteria’s for being a resource, although some capabilities may be considered as resources. Therefore in general, a capability is not a proper subcategory of the concept of Resource.

2.5 Capability Is Not Intentional

In some cases definitions of organisational capabilities include parts that refer to expectations, desires and intentions (expressions a and b below). However as we can see in (c) some capabilities that are not known but can be discovered, are not planned, emerge over time, not initially desired, and known but not part of the current value chain. Such capabilities may be of high value and a source of substantial revenues, as well as the basis for a new or changed core business. In (d) the existence of biological waste products can lead-to the production of biogas, which is not part of planned production.

> a) The ability to marshal resources to affect a predetermined outcome. [6]
>
> b) A manageable feature, faculty, function, process, service or discipline that represents an ability to perform something which yields an expected set of results and is capable of further advancement or development. [16]

c) The bakery realised that their cashier is capable of creating beautiful window displays that can be marketed and sold to neighbouring shops.

d) The bakery's waste products are capable to heat up houses through a waste management process.

Although plans, designs, wants, needs, expectations, desires and intentions are important, only some capabilities (and their results) are designed and intended.

From economical point of view capabilities with their intended results, as well as their side effects (undesirable results) play a role in decision-making, cost-benefit, trade-off, and scenario-analysis. Unrecognised capabilities may even be the cause to above average rent or sustained competitive advantage. Therefore in general, intentionality is not an essential characteristic of capability.

2.6 Capability Is Not a Function

Related to intention are teleology and the concept of 'function'. In expressions (a, b) we find two definitions of function; evolutionary and intentional design, and in (c) a third definition of business function.

a) "A function is a disposition that exists in virtue of the bearer's physical make-up and this physical make-up is something the bearer possesses because it came into being, either through evolution (in the case of natural biological entities) …

b) or through intentional design (in the case of artifacts), in order to realize processes of a certain sort." [17]

c) "A process or operation that is performed routinely to carry out a part of the mission of an organization". [18]

For all three variants of function we can find relevant examples of capability that are not functions. The example in expression (d) illustrates a capability case without intentional design, or an evolutionary physical makeup, or a routine organisational activity.

d) The bakery realised that their cashier is capable of creating beautiful window displays that can be marketed and sold to neighbouring shops.

Therefore in general, some capabilities are related to function but not all.

2.7 Capability Is Not Contextual

In the debate about the nature of capability the argument is sometimes made that all capabilities are context specific. This argument ties into philosophical, epistemological and ontological assumptions about an interlinked world, where knowledge is contextual and all propositions are true within a context, IST(context, proposition) [10, 19]. From a language and pragmatic point of view it is possible and also often desirable to discuss (and know) matter (see expressions (a, b, c, d) without a *full* set of details about contexts, situations and conditions.

a) Nils learned (what event/situation Nils need to be in) to bake a cake. [13]

b) Zlatan is capable to kick the ball with a speed over 100 km/h.

c) Zlatan is capable to kick the ball with a speed over 100 km/h, on a sunny day in September, in Malmö, and on a green field.

d) We want to acquire your capability to produce tasty cakes.

In the 'The Capability Approach' developed by the Nobel laureate Amritya Sen and in followup work [20], it is recognised that basic kinds of capabilities (expression e) such as natural, innate, and learned, are important.

e) "A newborn child has, in this sense, the capability for speech and language, the capability for love and gratitude, the capability for practical reason, the capacity for work"[20]

The question whether capability is (fully) contextual belongs partly to discussions about philosophical assumptions and contextualism. For pragmatic reasons, it is desirable to be able to define, know and reason about capabilities without explicit reference to context. An analogical argument against including context as essential characteristic is that definitions of resource, process, functions, service concept does not, in general, incorporate context as distinguishing characteristic. Therefore in general, context is not a distinguishing characteristic of capability.

2.8 Capability Is Not Abstract

Another frequently occurring argument is that a capability is essentially abstract, in the sense that it is not defined in terms of its properties and components, but by its use and the effects it can produce [21].

This argument has some problems with capabilities that are experientially, spatially and causally concrete. In the case of expression (a) the source part of the capability, my heart and veins, are known intimately, the lead-to process (pumping) is observable and the result, effect, outcome, oxygenated blood normally happens.

a) My hearts capability to pump and oxygenate blood.

It is possible the represent capabilities that are increasingly abstract. In expression (b) the source of the capability is specified as a category of persons. In (c) the lead-to process (create) and the result (comfortable home) are defined more abstract. In (d) the source is described by the fuzzy term jack-of-all-trades. In (e) the source is all together omitted from the capability representation.

b) The capability of a carpenter (a human with a hammer and nails), to cut and shape a functioning staircase.

c) The capability of a carpenter to create a comfortable home.

d) The capability of a jack-of-all-trades to facilitate a comfortable home.

e) The capability to facilitate a comfortable home.

Based on the cases (a-e) abstractness appears to be present in a *transferred* sense based on the constituent parts of capability. Another source of abstractness is that a capability represents is a *possibility* where the result may (normally) happen, based on some (disposed) source entities (in a possible world).

Although a capability represents a possibility there exist capabilities that range from very concrete to abstract, from a pragmatic perspective the abstract-concrete dichtonomy does not serve well as a distinguishing characteristic.

2.9 Capability Is Not a Process

The relation to the concept of 'process' is one of the most frequently debated topics. Is capability a new concept, or similar to process, or the same as process, or overlapping, are questions that are discussed. Unfortunately many discussions become language games where the term 'process' is a label for a great variety of meanings.

In order to avoid complicating the argumentation the term 'process' (or ‘basic process’) discussed here is taken from the Basic Formal Ontology (BFO) [17]. In BFO, 'process' is an occurrent (perdurant); an entity that unfolds itself in time or it is the instantaneous boundary of such an entity (for example a beginning or an ending) or it is a temporal or spatiotemporal region, which such an entity occupies. In a 'process' 'continuant' (material entities, humans, machines, etc.) entities participate.

In comparison with an example (a), from the business domain, the definition of the basic process is very raw, where the ‘business process’ is rich with additional qualified entities and underlying propositions. In expression (a) we find qualifications of process ('sequence of interdependent and linked procedures', ‘'convert inputs (data, material, parts, etc.)’, etc.), participating entities ('resources (employee time, energy, machines, money)', etc.), and participating results entities ('outputs', 'until a known goal or end result is reached').

> a) "Sequence of interdependent and linked procedures which, at every stage, consume one or more resources (employee time, energy, machines, money) to convert inputs (data, material, parts, etc.) into outputs. These outputs then serve as inputs for the next stage until a known goal or end result is reached." [22]

When breaking down the concept capability, in general three key parts appear in a pattern; a source, lead-to, and a result part (see following section for more information). In essence a substantial possibility that source entity(ies) lead-to a result". The source part is sometimes expressed as coordinated deployments of resources, combination of ways and means, perform activity, or ability, and the result part may be formulated as achieve outcome, desired effect, or address rapidly changing environments. The lead-to part represent the mechanisms (causal, formula, process, etc.) by which results are obtained based on the source and its elements. In expression (b) the source (team, skills, knowledge) lead-to (perform review process) to a results (successful completion of review process).

> b) The team is capable to perform reviews.

In order to fully discuss a capability all three parts must be considered. If any of them are left out then what is discussed in not a capability, although in many cases one part may be underdetermined or left out, and in other cases a part is qualified. In (c) the source (we) and the result (fulfil objectives) are defined and the lead-to mechanism is underdetermined.

c) We can fulfil our objectives.

Where does this leave process? Analysing the expression (a) we find additional semantics infused on top of a basic bfo:process, including intentional qualifications, and procedures, additions which reach beyond a general and commons sense conception of capability. The basic bfo:process can be associates with a lead-to mechanism that bring about some result, outcome, achievements, although a basic bfo:process does not by itself fill all slots in the capability pattern. We can also see that a basic bfo:process (and example (a)) does not include a possibility modality, normal occurrence or disposition, which is associated with capability.

Amongst practitioners a common practice can be found where the identification of, mapping out capabilities result in maps that looks like process maps [23]. Although a capability is not the same as a basic bfo:process, the capability of being capable of executing a (business) process or an activity (i.e. a processual ability) that lead-to a successful completion can be of interest and relevant to an organisation (see case (a)). However the added value of such capability maps may be difficult to explain to professional business process engineers.

An important pragmatic difference between processes and capabilities is that capabilities can, and in reality, transcend the boundaries of a single activity, routine, procedure, or well-defined business process. In the case of (d) the capability may be scattered across an organisation [24]. It is also common in capability definitions and approaches to add that entities, participating in a process, are integrated, organisationally embedded or combined in order to successfully lead to some results.

d) The capability to Set Right Price.

When shifting the focus to the result part, links to common practices, theories or paradigms such as Management by Result or Objectives, Program Logic [25] can be found. This, yet another link, indicate the uniqueness of the capability concept as a bridging concept that links together possessing, owning or having access to things or resources with a process the brings about some results or outcomes.

There are more to the concepts of capability than a basic, well-defined and encapsulated process. Therefor capability is not the same as basic process, although the basic bfo:process constitute an important part of a specific category of capabilities, the processual capabilities that lead-to a successful completion of the process.

2.10 Capability Is Not the Result

A key part of any expression of a capability is what comes after the 'lead-to', the results that are possible to bring about, normally happens, or are the result of a disposition.

Based on the three-part pattern (see next section for definitions) it follows that many combinations of source entities (person, material, process, etc.) can lead-to the same result. It also follows that a single combination of source elements can lead-to many different results. The separation of concerns between specification and realisation is used in frameworks such as UPDM[5], where the source entities are grouped into a Capability Configuration. It can be noted that authors often include a qualified source part in their definitions, that indicate that culture, coordination (a), integration, and combination of source elements of is of key importance.

> "a) repeatable patterns of action that are created through a firms management processes for coordinating its resource in processes for value creation." [4].

A vital consideration for any definition of capability is where the results are defined, measured along the capability chain/ladder, and applicable **capability horizon**, e.g. investment horizon, sustainable horizon, next quarter, time of exchange. The same bakery can consider different kinds of results (expressions b-f) along one or more lead-to chains/ladders, possibly including mediating entities, as relevant.

> b) A baker is capable to bake a cake.
>
> c) , to fulfil a management objectives
>
> d) , to make a cake that customers buys
>
> e) , to provide a tasty experience for the customer
>
> f) , to enable a sustainable bakery business

The result part is separated from a basic process and is not the only necessary element of the concepts of capability. Therefore capability is not the same as a result.

3 What Capability Is

Now that the capability construct has been stripped (deconstructed) of non-essential characteristics we shift focus to a demonstration of one general and common capability construct that covers the "what capability is not" aspects and distinctions. Furthermore this capability construct can be extended to create more specific capability concepts suitable for inclusion in subject fields, frameworks and theories.

Based on the preceding discussion and capability definitions found in dictionaries and existing capability literature five elements, in a relational pattern, can be identified. The identified elements are:

- a capability represents a *possibility* that some results can be brought into existence with respect to some source entities,
- *source entities* that participate in a thematic source role ('means', 'process', activity', 'combination of ways and means', 'realiser', capability configuration'), e.g. a baker, ingredients, baking skills,
- a relational and directional *lead-to* construct that provides a link between *source* and *result* entities ('to', 'by' 'generate', 'accomplish', 'achieve', 'that-makes', 'bring about', 'caused', 'realised by behaviour', 'normally happens'), e.g. baking process, enjoying process, speaking process,
- *result* entities that participate in a determinant thematic product role ('result', 'outcome', 'achievement', 'accomplishment'), e.g. entity, quality, trait, characteristic or functioning that comes into being, dies or does not come into being, a state-of-affair, a change, no change, something is performed, or done, or not done,

- a *substantiality* construct that indicate substantial, considerable qualities in participating entities (‘capacity’, ‘substantial freedom’, ‘sufficient’, ‘adequate’, ‘requisite’, ‘feasible’, ‘disposition’).

In essence - “a substantial possibility that source entity(ies) lead to a result".

The following diagram illustrates a conceptual schema for the capability concept and its five key elements.

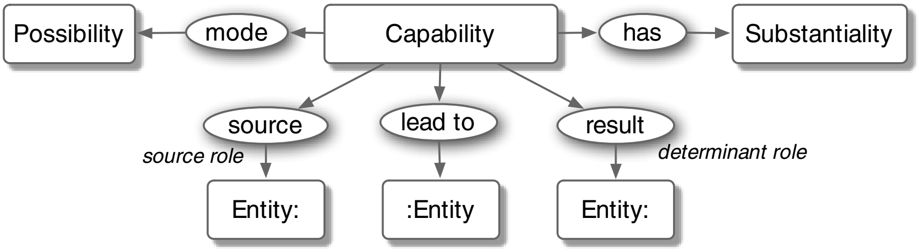


Fig. 1. Capability conceptual graph

The following table demonstrate an interpretation of expression (a) based on the five elements.

a) The baker has the capability of baking a cake that customers praise for its taste.

Table 2. Key elements expressed in (a)

Capability Element	**Interpretation**
Possibility	The baker can bake a cake if the baker choose to do so
Substantiality	The baker has good ingredients, suitable oven, skills to bake tasty cakes, and has substantial freedom to bake
Source	Baker, Ingredients, Oven accessible by Baker
Lead-to	Baking process, selling process, tasting process, praising process
Instrumental Result 1	Baked cake
Instrumental Result 2	Sold cake
Terminal Result 3	A cake that customers praise for its taste

As a second demonstration the capability definition from UPDM v2.0 [26] in expression (b) is interpreted.

b) "The ability to achieve a desired effect under specified [performance] standards and conditions through combinations of ways and means [activities and resources] to perform a set of activities"

Table 3. Interpretation of example (b) and key elements

Capability Element	Interpretation
Possibility	by the combinations of ways and means a desired result can be achieved
Substance	- *(not explicitly stated in definition)*
Source	resources (material, performer, information, …), ways and means, skills to perform activities, ability
Source modifier	guided and constrained by specified [performance] standards
Source modifier	activities are performed under specified conditions
Lead-to	Achieve, by performing a set of activities
Lead-to modifier	combinations of ways and means [activities and resources]
Lead-to modifier	guided and constrained by specified [performance] standards
Lead-to modifier	activities are performed under specified conditions
Result	effect (a state of resource)
Result modifier	effect is desired by someone

As a third demonstration, examples of how the discussed aspects and distinctions may be incorporated in more specific capability constructs that are relevant to specific subject fields, situations, theories, or frameworks.

Table 4. Examples of how aspects and distinctions may be incorporated in specific constructs

Aspect, Distinction	Possible incorporation
What-How	*Not applicable to the definition of the concept of capability*
Organisation	State that a specific organisational capability only considers source, lead-to, and result entities that are attributed or relevant to an organisation.
Single kind	Persons involved in different kinds of work perspectives are often interested in different kinds of capabilities, which involve specific types of source, lead-to, and/or result entities, i.e. define a type of capabilities for a work perspective.
Resource	State that only valuable, controlled, or owned source, lead-to, and/or result entities are relevant, e.g. incorporate Resource in the definition.
Intention	Modify or qualify source, lead-to, and/or result entities with intentional propositions, e.g. desired result.
Abstract	State that only abstract source, lead-to, and/or result entities are relevant.
Function	State that only source entities that has a function are relevant
Context	Add context specific modifiers to source, lead-to, and/or result entities. OR Add a Is –True-in-Context (IST) operator to the framework or theory.
Process	Modify source, lead-to, and/or result entities to define a Processual Capability that match the structure of a process definition, e.g. the lead-to process has input and output source entities, and a result entity - is fulfilment of goals.
Result	A result is always present for a capability and can be modified.

More specific capability variants can be constructed based on a general capability concepts using John Sowa's Lattice of Theories approach [27], and its propositional operators - expand, contract, revise.

By enabling authors to openly expand a general capability concepts by adding and motivating propositions, assumptions, contexts, constraints, modifiers and qualifiers, it is suggested that the process of integrating (in theories and frameworks), comparing and relating capability definitions becomes easier. This is furthermore supported by the use of the practice of naming specific entities differently than general entities.

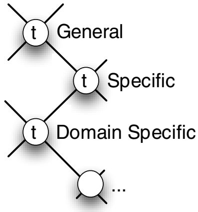


Fig. 2. Illustration of specific concepts defined by adding propositions using the expand operator according to the Lattice of Theories [27] principles

It is worth noting that many capability definitions include a *dynamic* part that links source entities to the results, e.g. "combinations of ways and means". Such dynamic aspects often separate the perspective capabilities offer from more static views of an organisation, such a viewing an organisation as a set of owned or accessible resources. The *lead-to* part provides a straightforward basis for testable hypothesis, such as if explicit combination of ways and means is better than no combination.

The demonstrated capability construct indicates that a general and common sense conception of capability is possible to find, which is relational and unique in its support of all discussed "is not" aspects and distinctions.

This capability construct can also be identified in adjacent concepts, such as a business process that fulfils objectives; a function has purpose, servicing (use of effort) that lead to benefit of another, and resource lead-to a value. Given the underlying pattern it is not surprising that capability is perceived as being the same as, or overlapping, or more specific than adjacent concepts.

4 Conclusions

In this paper we started out with the observation of fragmented capability definitions, theories, and approaches, and asked the question if the advancement of broader agreements could be benefited from understanding of what capability is not.

This paper and example cases cover the surface of various subject fields and possible applications of a capability construct. Despite the limited set of covered aspects, distinctions and applications, the identified example cases, and argumentation provides challenges to many existing capability constructs in the sense that these constructs are not likely to be considered as 'the' definite one-and-only and correct construct, applicable in all cases. The argumentation, related to aspects such as abstract-concrete and related concepts such as resource, suggests that many capability constructs constitute specific variants of a more general and common sense capability concept.

Through the use of a deconstruction and reconstruction technique authors are enabled to openly declare their (domain) specific additions assumptions, situations, propositions, constraints and qualifications) to a general and common sense semantics, and together with proper labelling, naming conventions this promise to greatly simplify integration (into theories and frameworks) and comparisons of capability constructs.

Based on the cases and argumentation I argue that an approach of identifying "is not" aspects and distinctions will enrich discussions, definitional work and harmonization efforts amongst researchers and practitioners in various subject fields.

References

1. Longman Dictionary of Contemporary English Online
2. Ogden Basic English, http://ogden.basic-english.org/
3. Dosi, G., Nelson, R.R., Winter, S.G.: Nature and Dynamics of Organizational Capabilities, pp. 1–404 (2008)
4. Sanchez, R.: A scientific critique of the resource-base view (RBV) in strategy theory, with competence-based remedies for the RBV's conceptual deficiencies and logic problems. Elsevier (2008)
5. Group, O.M., Inc: UnifiedProfile for DoDAF and MODAF (UPDM) (2012)
6. McKeen, J.D., Smith, H.: IT Strategy, 2nd edn. Prentice Hall (2011)
7. OASIS: Reference Model for Service Oriented Architecture 1.0. OASIS (2006)
8. Robeyns, I.: The Capability Approach: An Interdisciplinary Introduction, pp. 1–57 (2003)
9. ISO, IEC, IEEE: ISO/IEC 42010 Systems and software engineering—Architecture description (2011)
10. Sutton, R.I., Staw, B.M.: Whattheory is not. Administrative Science Quarterly, 371–384 (1995)
11. Baskerville, R.: What design science is not. Eur. J. Inf. Syst. 17, 441–443 (2008)
12. Ulrich, W.: Defining the Business Capability - A CheatSheet, http://www.bainstitute.org/resources/articles/defining-business-capability-cheat-sheet
13. Ginzburg, J.: How to resolve how to. In: Bengson, J., Moffet, M.A. (eds.) Knowing How: Essays on Knowledge, Mind, and Action, pp. 1–36. Oxford University Press (2011)
14. Prahalad, C.K., Hamel, G.: The core competence of the corporation. Harward Business Review
15. Cardeal, N., António, N.: Valuable, rare, inimitable resources and organization (VRIO) resources or valuable, rare, inimitable resources (VRI) capabilities: What leads to competitive advantage? Afr. J. Bus. Manage. 6 (2012)
16. Information Technology (IT) Discipline Management Capability, http://www.if4it.com/
17. Smith, B., Grenon, P.: Basic Formal Ontology (BFO), http://purl.obolibrary.org/obo/bfo
18. Business Function, http://www.businessdictionary.com/definition/business-function.html
19. McCarthy, J.: A Logical AI Approach to Context (1996)
20. Nussbaum, M.C.: Women and Human Development. Cambridge University Press, Cambridge (2009)
21. Bergholtz, M., Johannesson, P., Andersson, B.: Towards a model of services based on co-creation, abstraction and restriction. Presented at the 6th International Workshop on Value Modeling and Business Ontology (2012)
22. Process, http://www.businessdictionary.com/definition/process.html
23. Harmon, P.: Capabiliites and Processes, http://www.bptrends.com/publicationfiles/advisor20110712.pdf
24. Dutta, S., Zbaracki, M.J., Bergen, M.: Pricing process as a capability: a resource-basedperspective. Strat. Mgmt. J. 24, 615–630 (2003)
25. McCawley, P.F.: The Logic Model for Program Planning and Evaluation, pp. 1–5 (2002)
26. Ontologies and terminologies: Continuum or dichotomy? 1–12 (2012)
27. Sowa, J.F.: Language Games: A Foundation for Semantics and Ontology. Game Theory and Linguistic Meaning, edited by Ahti-Veikko Pietarinen, pp. 17–37. Elsevier (2007)

Analysis of Agile Software Development from the Knowledge Transformation Perspective

Ilia Bider

Department of Computer and Systems Sciences (DSV),
Stockholm University (SU)/IbisSoft AB, Stockholm, Sweden
ilia@dsv.su.se, ilia@ibissoft.se

Abstract. While the Agile Software Development (ASD) has been successfully promoted in the last 15 years, there is no agreement on how to determine whether a particular project is agile or not. Some practitioners consider agility as strict usage of a specific methodology, e.g. SCRUM, others consider agility as adhering to Agile Manifesto. The lack of common view on ASD prevents creating common guidelines on when the usage of ASD is appropriate. This paper presents a model of ASD that helps to differentiate it from the traditional, phase-based development, and more strictly defines the area of its applicability. The model has been built based on the knowledge transformation perspective, as the author considers it to be the most differentiating perspective when comparing ASD to traditional software development. For building the model, the ideas from SECI model of Nonaka have been exploited. The results, in the form of requirements to be fulfilled for successful employment of ASD, are demonstrated through analysis of completed ASD projects.

Keywords: Software Engineering, system, software project, agile, knowledge transformation, SECI model.

1 Introduction

Agile Software Development (ASD) has appeared as a reaction on the increasing rate of changes in system requirements, see, for example, [1]: "requirements change at rates that swamp traditional methods". Though there are numerous books and articles on ASD, e.g. [2], there is no complete agreement on the definition of what ASD is. The main problem here is that the essence of ASD is sometimes mixed with particular project methodologies, like SCRUM, that could be used in both agile and non-agile software projects.

The existence of the confusion of what agile means could be easily seen when investigating the online discussions devoted to the topic, such as "Do you agree or disagree that Scrum is not Agile?" [3]. This corresponds to the authors own experience of investigating software projects referred to as agile because they use SCRUM, despite the fact that they concern the development of hardware-near software with for each release fixed requirements specifications [4].

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 143–157, 2014.

Currently, the only definition of ASD on which there is a common agreement, is the so-called agile manifesto [5]. The weak side of the manifesto is that it does not have an underlying scientific theory, but consists of a number of principles that allows different interpretations. For example, [6,7] consider the lack of theory and philosophy behind Agile Manifesto as a root of misunderstanding of agility not only in software development, but also in other related disciplines, e.g., Information Systems (IS). Their consideration is a result of reviewing agility in many disciplines. They also claim that the lack of considering agility outside the software development makes this manifesto not generally accepted by the wider audience.

The goal of this paper is twofold, namely, it is aimed to suggest a theoretical framework that could help in (a) explaining the difference between the agile and non-agile software development; (b) defining the area of applicability for ASD. In this paper, we are not focused at discovering new facts about ASD, but rather on finding a practically useful framework to explain the existing facts, so that they can be understood by all stakeholders concerned, including management of software vendors and customers.

To attain the goal above, we will build a dynamic model of a software project that employs the ASD methodology and compare it with the same kind of model for traditional, phase based software development. Based on this model, and comparison with traditional development, we will derive the requirements on the software project in which ASD could be successfully employed.

When building a model for the purpose of analysis of agile development, we took a knowledge transformation perspective on software development. When developing this perspective, we used the ideas of Nonaka [8] on knowledge transformation in organizations. In particular, our model was influenced by Nonaka's SECI model, where SECI stays for *Socialization – Externalization – Combination – Internalization*. The SECI model explains how the knowledge is transferred/and or converted inside and between two different forms – tacit (in the heads of people) and explicit (in the form of writing, drawings, mathematical models, etc.).

The paper is structured in the following way. In Section 2, we give a short overview of our method and background on which we built our framework. In Section 3, we present an informal description of the proposed framework. In Section 4, we use this framework for creating requirements on the project that need to be fulfilled in order for the agile approach to be successfully employed. In Section 5, we apply the results from Section 4 to analysis of past software projects from the literature and from the author's own practice. In Section 6, we discuss the result achieved so far, and outline plans for the future.

2 Method and Background

In our research, we employ a Design Science Research (DSR) approach, more specifically its interpretation according to [9]. In this interpretation, DSR, as a way of generating and testing hypotheses for generic solutions, requires researchers to act in two different worlds: (a) the real world of specific problems and solutions in local practices, and (b) the abstract world of generic situation, problems and solutions. There is no specific requirement on the order in which the movement is completed. Researchers can start with searching a solution for a known in a local practice problem before or after generalizing it, or they can start with building a solution for the problem unknown, and then finding what the solution is good for. The main point is to have in the end a description of the triad <generic-situation, generic problem, generic solution> and one or several test cases that shows that the generic solution applied in a specific situation can solve the problem.

In our case, we are identifying conditions for success when the agile approach to Software Engineering (SE) is employed in the project. These conditions can be used for solving a number of problems that exist in SE practice that are discussed later in Section 4. In section 5, we will present test cases of using the solution developed for analysis of completed projects.

Agile system development is a concept that includes many variations, like eXtreme programming, test-based development, peer programming, etc. To answer the question of agile applicability in general terms, we cannot rely on building a model based on a particular variation/interpretation of Agile Software Development (ASD). The model we need should be on a higher abstract level that includes no details, but all essential features of the agile approach. Our model of ASD and explanation of how it differs from the traditional one was built based on the following four sources:

1. Own experience of software development, agile as well as not agile, in various capacities, e.g., as a programmer and a group leader in larger projects, and as a project leader and architect in smaller R&D projects, see for example, [10].
2. SECI model of Nonaka [8] on knowledge transformation, which gave us the idea to build a model based on knowledge transformation perspective. More exactly to consider software development cycle as knowledge transformation cycle.
3. Good regulator theorem of Conant and Ashby, which gave us the idea to introduce a new type of knowledge absent from SECI model – embedded knowledge.
4. Agile Manifesto [5]. We have chosen Agile Manifesto from the vast literature on ASD based on the following reasons: (a) it comprises the essence of ASD without going into details of specific methods; (b) as far as we know, this is the only document on which the community of practice agrees, therefore it can be considered as a collective wisdom of the community.

We do not claim that the sources above are the only ones that could be used for answering the question posed in the title. However, they were sufficient for us to design the solution we were looking for. Other sources, when discovered can be used for validating and refining the solution, some of them are mentioned in Section 6.

As sources 2 and 3 are less known in the domain of SE, below, we present a short overview of them. The SECI model, where SECI stays for *Socialization – Externalization – Combination – Internalization*, by Nonaka [8] explains the ways of how knowledge is created in an organization while being transformed from the tacit form (in the heads of the people) to the explicit one (e.g. on the paper in the form of texts, drawings, etc.) and back, see Fig. 1. [8] defines the cycle of knowledge creation as consisting of four steps or phases:

1. The cycle starts with *Socialization* (left top corner of Fig. 1), where tacit knowledge is transferred from the heads of one group of people to others via

informal means, conversations during the coffee breaks, meetings, observations, working together, etc.

2. The next phase is *Externalization*, which is the conversion of knowledge from the tacit form into the explicit one, e.g. a model of situation (right top corner of Fig. 1).
3. The third phase is *Combination*, which is transforming the externalized (explicit) knowledge in a new form using existing knowledge, e.g. solution design principles (right bottom corner of Fig. 1).
4. The last phase is *Internalization*, which converting the explicit knowledge, e.g. a solution, in the tacit knowledge of people that are ready to apply this knowledge to any situation that warrants it (left bottom corner of Fig. 1).

The cycle of SECI can be repeated indefinitely reflecting constant creation of new knowledge.

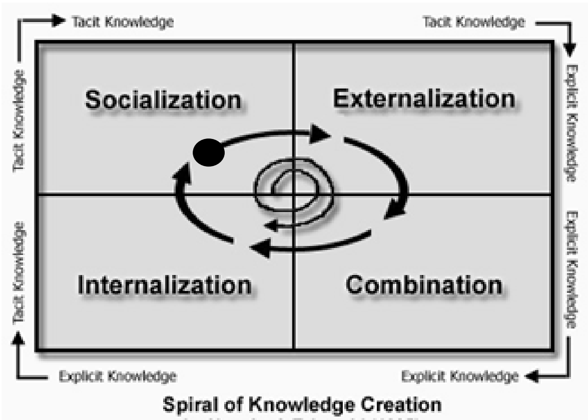


Fig. 1. SECI diagram, adapted from [8]

The SECI model identifies two type of knowledge tacit and explicit. This is not enough if we want to apply knowledge transformation perspective to software development. We need to consider the knowledge built in the software system, creating which is the goal of any new software development project. This knowledge cannot be regarded as explicit as people using the system may not know how it works internally. This knowledge cannot be called tacit as it belongs to the system not to the human beings (unless we consider the system being an agent comparable with the human being). We refer to the knowledge built-in into the system as to embedded knowledge. The existence of embedded knowledge is indirectly supported by Conant and Ashby theorem [11] that "every good regulator of a system must be a model of that system". Using analogy, we can reformulate (specialize) the Conant and Ashby theorem as:

"*Every good software system is a model of the requirements its implements/satisfy*"

The latter is indirectly conformed by the practice of reengineering of legacy systems for which the source code has been lost or is so dirty that it requires too much time to understand the details. It also is in line with the author's own experience of supporting and further developing legacy system built by unknown people.

Adding an additional category of knowledge, we applied the ideas from [8] to the cycle of software development both in its traditional and agile forms, and built models that represent these methodologies of software development. In this paper, we limit ourselves to discussing full cycle of new product (software) development, leaving the issues of applicability of the agile methodology to maintenance of the existing products built in the traditional way outside the scope of this paper.

When considering requirements on software project for successful employment of ASD, we take a system perspective on the software project and its environment. More exactly, we consider that there are three distinct but interconnected systems involved in any software project:

1. The software system (S-system), i.e., the virtual artifact being developed or modified.
2. The software project (P-system), i.e., the work system undertaking the development or modification of S-system.
3. The software context (C-system), i.e., the environment in which the software product is being, or is intended to be used.

The three-system division introduced above relates to many existing studies and frameworks. For example, [12] distinguished between the Narrow System of Interest (S) and Wider System of Interest (P), and argued that both exist in an Environment (C) and even Wider Environment. We are not going to discuss the system view any farther, but will use the division into S, P, and C systems when discussing requirements, as different requirements will concern different systems, or interconnection between different systems.

3 Building Models of Software Development

3.1 A Model of Traditional Phase Based Software Development

Applying Nonaka's ideas to knowledge transformation during the traditional cycle of software development, we have got a model presented in Fig. 2. The model is abbreviated to ECEA, which stays for *Externalization-Combination-Embedment-Adoption*. The cycle starts with tacit knowledge on problems and needs in the heads of the stakeholders (human participants of C-system) and consists of four phases:

Phase 1 – Requirements engineering (RE) - consists of transforming tacit knowledge possessed by C-system participants into explicit knowledge of requirements (right top corner of Fig. 2). This phase corresponds to *Externalization* from SECI model. The conversion is usually done via facilitating workshops where specialists in requirements engineering work together with the stakeholders and, possibly specialists in the domain of C-system. The stakeholders and domain specialists have tacit knowledge on the C-system with its problems and needs, while the RE specialists help to convert this knowledge into a structured description in form of requirements.

Phase 2 – Design – consists of transforming explicit knowledge of requirements to explicit knowledge software system design (right bottom corner of Fig. 2). This phase corresponds to *Combination* of SECI model. The conversion is done by applying contemporary or innovative design principles to the requirements in order to produce design documents, e.g. class diagrams, user cases which can be at the next phase converted to a software system.

Phase 3 – Coding consists of transforming the explicit knowledge of system design to the knowledge embedded in the system (left bottom corner in Fig. 2). Though parallel to *Internalization* from SECI model, this phase does not correspond to the latter exactly. We call this transformation *Embedment*. For programmers engaged in this activity, this transformation is from one form of explicit knowledge to another (program code). However, for the C-system, the knowledge built into the S-system is not explicit. The conversion is done using contemporary or innovative methods of programming, e.g. object-oriented programming.

Phase 4 – Learning to use in own practice – consists of transforming embedded in the system knowledge into the tacit knowledge of how to use it (left top corner in Fig. 2) in C-system. Though parallel to *Socialization* from Fig. 1, this phase does not correspond to the latter exactly. We call this transformation *Adoption*. If the system is aimed to work in symbiosis with people "inhabiting" C-system, adoption includes the users learn how to use the system when completing their tasks. In case of pure technical systems, adoption includes adjusting other component in C so that they can successfully work together. Normally, both types of adoption happen for any S-system. Even when the system is purely technical, some human participants of C-system need to have tacit knowledge on the relationships between the new S-system with other components of the context system. The latter is important in case a problem in C related to S arises.

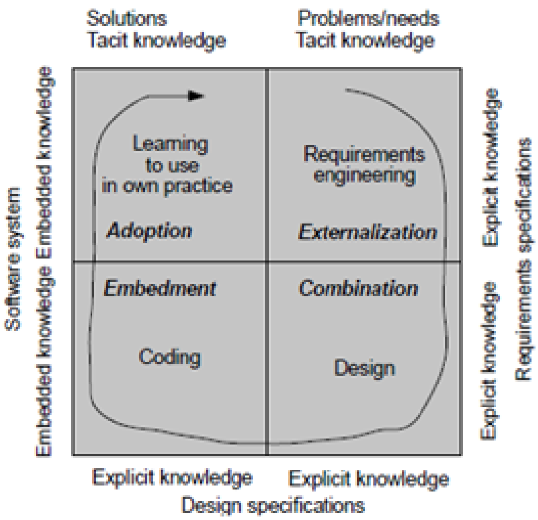


Fig. 2. ECEA - Knowledge transformation in the traditional process development

Note that the cycle above is idealization. For example, parallel to *Embedment* and *Adoption* there might be additional knowledge transformations:

- Writing manuals, which is converting embedded knowledge into the explicit one,
- Reading manuals, which is converting explicit knowledge in manuals into the tacit one

However, these two transformations are losing its importance as the contemporary users learn new systems in the try and error fashion without bothering to read manuals.

As with the SECI model, the cycle of Fig. 2 can be repeated as soon as there are needs to change the software system. We also need to point out that the phases represented in Fig. 2 need not be executed in a *strict sequential order* as prescribed by the waterfall methodology. The first three phases can run in parallel and affect each other. For example, the RE may pass the main requirements to the designers without uncovering all requirements, so that the designers can analyze whether the requirements already discovered can be implemented in a software system, or some of them need to be renegotiated. The same kind of *parallelism* can be arranged between *Design* and *Coding*. This comment means that the model on Fig. 2 is not representing the waterfall model only, and it allows iterations and conversations between the phases, as it is done in so-called V-model. These details, however, are not essential for the purpose of the modeling accepted in this paper.

Summarizing the deliberation above, as traditional we consider software development in which all four phases are *separated*, independently whether they are completed in *sequential* (e.g., waterfall model) or *non-sequential* (e.g., V-model) order. This is different from the agile development as will be shown in the next subsection.

3.2 A Model of Agile Software Development

From own experience, as well as from literature on Agile System Development (ASD), see for example [1,13], the main difference between Traditional System Development (TSD) and ASD is that the latter relies much more on the tacit knowledge than on the explicit knowledge. In terms of the knowledge transformation cycle of Fig. 2, ASD tries to shortcut the phases of *Externalization* and *Combination* and go directly to *Embedment.*

Shortcutting *Externalization* and *Combination* in Fig. 2, we get a model in Fig. 3. The model is abbreviated to SEA, which stays for *Socialization-Embedment-Adoption*. The SEA cycle differs from ECEA cycle in three respects:

1. The nature of the first phase in Fig. 3 is changed against Fig. 2. It consists in transferring tacit knowledge on the problem and needs from the stakeholders to the development team. This phase corresponds to *Socialization* in Fig. 1.
2. Design and coding are merged into one phase *Embedment*, which we continue to call *Coding*.
3. In addition, one big cycle is substituted by many smaller and shorter ones. The system is built iteratively starting with the basic functionality. During the exploitation of the basic system, better understanding of the needs is acquired, which is converted in adding details to the system in the next iterations. In other words, agile methodology is based on a simple motto “Develop and introduce in practice as little as possible as soon as possible, and build upon it in the following iterations”.

As with ECEA, the SEA model is an idealization. During both *Socialization* and *Embedment*, some knowledge is explicated, e.g., simple flowcharts drawn on the paper, or on the black/white board, or user stories are written as plain text. This, however, is done in order to facilitate the tacit knowledge transfer from one group of people to another, or into the system, not to make all knowledge explicit. Therefore, using explicit knowledge in *Socialization* of Fig. 3 is different from its usage in *Externalization* of Fig. 2, where the whole phase is directed at creating explicit requirements before going further with *Combination* and *Embedment*.

As with the traditional development of Fig. 2, having different phases in agile development does not mean that they *need to be implemented in sequence*. Both *Socialization* (phase 1), and *Embedment* (phase 2) can be completed *in parallel*.

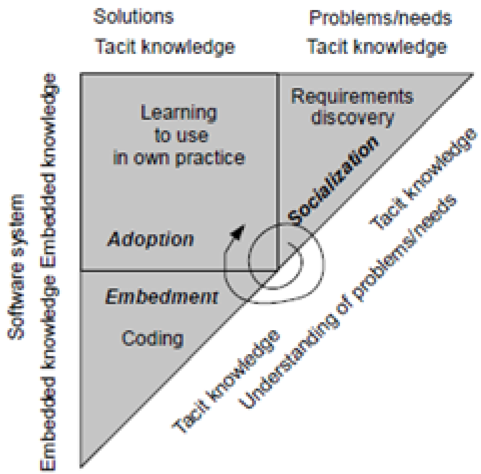


Fig. 3. SEA - Knowledge transformation in the agile software development

4 Analyzing the Models

4.1 Advantages and Drawbacks of Traditional Development

In this section, we will use the ECEA model to analyze strengths and weaknesses of the traditional approaches to software development. Base on phases identified in Fig. 2, we can identify the following advantages of the traditional development.

1. There are 4 distinct phases of knowledge transformation that allows distributing the work between the experts in four different areas: *RE*, *System Design*, *Programming*, and *Adoption* (including user training, writing manual and project management for system adoption). Using explicit form as an output of *Externalization* and *Combination* facilitate communication of the specialized experts in different areas.
2. Having explicit requirement specifications, and design specifications facilitates using existing knowledge on design principles and programming.

3. Having explicit requirement specifications allows entering a contract agreement with fixed obligations on the side of the development team.

The drawbacks of using traditional development are the other side of its advantages:

1. *Instability*. A small error or missing a requirement in *Externalization* may be amplified in other phases. As connection to the tacit knowledge of the problem and needs is absent in *Combination* and *Embedment* there is little chance that a miss will be recovered at these phases. In addition, specialization may result in the experts in the design and programming having no knowledge, skills or experience to communicate with the stakeholders.
2. *Fuzzy requirements*. Using requirements as a basis for design, and a contract agreement presumes that all such requirements could be understood and formulated, and the customer has skills and experience of imagining how the C-system will work when an S-system that satisfies the requirements is put into operation. In situations when there is a reference point, for example, an older S-system already exists in the C-system, such a presumption could be warranted. However, when such reference point does not exist, for example, when the C-system should be significantly changed when a new S-system is deployed, such presumption is not warranted. In the latter case, insisting on fulfillment of the unchangeable requirements as a basis for the design and the contract could lead to a catastrophe, more exactly, the system satisfying all requirements, but being useless.
3. *Evolving context*. For a large complex S-system, it may take too much time to go through one big cycle. While a new S-system is being developed, the C-system can undergo changes, e.g., problems and needs may continue to evolve. As the result, the S-system could be outdated before it is deployed in practice of C-system. This could become a major obstacle for dynamic contexts.

4.2 Requirements on Agile Software Development

Agile software development is a reaction on the drawbacks of the traditional one. Its main principles are directed toward mitigating the three problems (drawbacks) listed in Section 4.1, namely:

Socialization as the first phase of the agile cycle of Fig. 3 is aimed at mitigating the first drawback of STD (see Section 4.1). It ensures that the whole development team gets tacit knowledge on the problems and needs that they convert into the S-system at the *Embedment* phase. Running *Socialization* and *Embedment* in parallel allows the team to refine their understanding of problems and needs even after starting writing the code.

Focusing on producing software without producing requirements and design documentation is aimed to mitigate the second problem on the list of drawbacks of TSD from Section 4.1. Demonstration of software can spin imagination and facilitate refining understanding of the needs based on the possibilities to be provided by S-system.

Splitting one big cycle into a series of smaller ones is aimed at mitigating the third problem on the list. It allows putting into operation the first version of the S-system in a shorter time, and thus avoiding it being outdated for a dynamic C-system. Running *Socialization*, and *Embedment* in parallel also contributes to the same goal as changes in C-system can be discovered and take care off on the later phases of the development. Splitting the cycle also helps alleviating the second problem, as only actual introduction of an S-system into the context (C-system) can show how well these two systems fits each other. A misfit of the first version is possible to fix in the next version.

However, while alleviating the problems of the traditional software development, agile model sets requirements on all three interconnected system P, S and C, thus limiting its area of application. Below we list the requirements based on the analysis of SEA model of Fig. 3:

1. As there is no explicit requirement specification or design documents, the whole development team should be engaged in acquiring tacit knowledge from stakeholders. It means the needs of switching from specialization of traditional team to universality of the agile team. All members need to have knowledge on all aspects of software development, from business and system analysis to programming. It does not mean that all of them need to be equally proficient in all branches, but they need to have enough understanding of all of them to be able to acquire and transform the knowledge through informal meetings without writing lengthy documents. This is a requirement on P-system - you cannot run an agile project if you do not have such a team or cannot create it in the first phase of the project.
2. Equally important is the attitude of the stakeholders inhabiting C-system. Firstly, they need to assign time to work with the P-system team during the whole project, not just in a short series of facilitating RE workshops. Equally important is that they have patience when communicating with the developers who might not know their domain and can raise seemingly silly questions.
3. As there are no explicit requirements there is no standard document on which a contract between stakeholders and developers can be based. The relationships between them should be regulated having some other scheme in mind.

One of the main features of the agile software development is that instead of delivering an S-system in one step, the delivery is split in several cycles each of them resulting in delivering an enhanced S-system for its implementation inside C-system. Each cycle should be finished in the shortest possible time to mitigate the problem that arises from the dynamic context. The following three requirements are connected to the cyclic delivery:

4. Providing the first delivery in a short time implies that a relatively small core S-system could be successfully implemented in the given C-system. Actually, this is a requirement on C-system, which can be expressed as possibility to identify a small core S-system implementation of which has sense (is meaningful). This, obviously, is not always true. For example, suppose a new S-system should substitute an older (legacy) system. It is difficult to imagine that a new core system that lacks some essential functions of the old one will be acceptable for such the substitution.
5. Splitting the delivery in a number of smaller cycles does not exclude, but rather promotes the final system becoming quite complex. Providing short delivery time for each iteration requires that enhancements added to each release do not require totally rewriting the code of the core system or the code produced in previous iterations. Rewriting will increase time needed to produce any meaningful enhanced

version of the system, and this time will increase with each iteration. Avoiding re-writing the code requires the core S-system having solid architecture that ensures the code remaining more or less the same during the whole project (lifetime of P-system). This is actually a requirement on S-system.

6. Short delivery time for each iteration requires high efficiency of coding which puts requirements on the languages and tools used for developing an S-system. It is not practical to use low-level languages; an agile project needs to employ high-level means that increase the productivity. These can be high-level domain specific languages, like 4th GL languages of 1980th and 1990th (that disappeared as class), libraries of functions or objects, development platforms, like Ruby-On-Rails, or executable models. These means could be acquired from a third party from the beginning of the project, or developed in parallel with the S-system produced by the project. These means should be synchronized/aligned with the overall architecture of S-system. The requirement of employing high-level means can be considered as a requirement on P-system, and partly on S-system, the chosen means need to be aligned with the S-system architecture.

The list of requirements above can be used for several purposes, including:

1. *Analyzing* what went wrong/right in a successful or unsuccessful attempt of employing agile methodology.
2. As a *check list* for decision whether employing agile methodology have chances for success. For example, if a customer cannot be convinced to ease the demand of having a contract based on the detailed requirements documents, employing agile methodology could be a risk factor. The same is true if there is no possibility to agree on a relatively small core S-system for the first iteration. Having a team with highly specialized members may also prevent agile development. In this case it might be preferable to use phase based scheme of Fig. 2, but run the phases *in parallel*.
3. As part of the *plan of action* when decision to use the agile approach has been made. For example, if agreement on the core S-system and agile methodology has been reached, one could hire a proper team, and choose proper means to boost productivity and ensure stable architecture.
4. The requirements above plus the models in Fig. 2 and 3 can be used as *educational materia*l that clearly shows the difference between the traditional and agile methodologies, weakness and strength of both, as well as their areas of applicability. Considering that the models are independent from a particular brand of traditional or agile methodology, such educational material has a potential to be widely used.

5 Analysis of Success and Failures of ASD Projects

To make the first verification of the suggested in Section 4.2 requirements, we analyzed two cases of completed ASD projects (item one on the list of the possible usage of requirements). The first case concerned an unsuccessful ASD project we found in the literature. The second case was from our own practice; it concerned the development of the system described in [14]. Due to the lack of space, in this paper, we can only give a short overview of the results of the analysis of these two projects.

5.1 An Attempt of Substituting a Legacy System in an Agile Manner

The story of this project was presented at Agile Conference in 2008 [15] under the name "When Working Software Is Not Enough: A Story of Project Failure", and then repeated at other events [16]. This project has been chosen for analysis due to the following reasons. Firstly, it is relatively well documented: besides a short presentation of the talk [15], the video record of the talk [17], and slides of the presentation [16] are available on the web. Secondly, this documentation contains analysis of the reasons for failure that can be compared with our own analysis. Thirdly, in addition to the analysis of the failure given by the project manager [15,17,16], an analysis by an independent expert is also available [18].

Based on the description [14,15], video recording [17], and independent analysis [18], we can conclude that the main reason for the project failure identified by the project manager corresponds to not fulfilling our requirements 2, and 3, i.e. stakeholders doing their part of job, and proper relationships between the people inhabiting C and P systems. As far as requirements 1 is concerned, the project manager considered the team being mature, and the materials analyzed does not provide enough evidence to make an independent judgment.

As far as requirements 4 is concerned, identifying a smaller core S-system that could be implemented in practice, the project failed to do so. They interpret the motto "Deliver working software frequently" only as demonstration, and testing by the customer (which has not been done as follows from the video recording [17]). The project could not identify and agree on a small core S-system to be directly introduced in practice. Though this fact is revealed in the presentation, especially in the video recording, the analysis did not consider failure to agree on a core system as a major reason for the project failure.

As far as requirements 5 and 6 are concerned, it seems that they were not fulfilled. The coding was done in C# and SQL, and changes in the requirements priorities caused a lot of code rewriting. The latter points out to lack of solid architecture. What is more, the analysis in [15,17] does not considers solid architecture and means of enhancing productivity to be prerequisites for the agile project success.

Summarizing the above, the analysis of this project based on the requirements from Section 4 confirms the analysis of the project manager as far as requirements 2 and 3 are concerned. In addition, it highlights the project not fulfilling the requirements 3 (identifying a smaller core S-system), and 5 and 6 - lack of solid architecture and means that enhance the efficiency. In our view, the last two contributed to the project failure in no less degree than not fulfilling other requirements.

5.2 Developing a System from Scratch in an Agile Manner

The system called ePortal was developed for a large Swedish call center to solve the daily staffing problems that can be defined as follows. The system was developed by IbisSoft AB in the agile manner in tight collaboration with her customer Eniro AB, the author serving as a team leader of the IbisSoft's team. The core system was successfully developed and introduced in organizational practice in the spring of 2009. The project went through three major cycles of which the second was the most critical as it include adding 600 new users (agents) that could work from both home and the office. The system is fully operational and continues to be used without major problems.

Considering the requirement from Section 6, the project satisfied the requirements 1-6 in the following manner:

1. The development team consisted mostly of two members (the project leader including). Other specialists were called for solving or investigating specific technical problems. The team had several years' experience of working together without detailed requirements or design specifications using face-to-face communication as the main channel for communication.
2. The customer found it very convenient to work in a relatively informal manner around first a prototype, and then - emerging system. The customer was a geographically distributed organization. Only IT function was in the same city as the development team, other stakeholders (including system users) were spread all over the country. Face-to-face meetings between the developers and stakeholders were rare. Communication was managed through (a) email, (b) phone, (c) the system under development that was accessible for the stakeholders all the time. All testing was done by the customer, and the customer was responsible for implementing the system in their organization. The implementation was well planned, including the customer educational department creating a video on how to use the system.
3. As the project did not rely on detailed requirements, the contract was based on the general estimation of the amount of work needed. Prioritization was done on the fly moving some features to the future releases or adjusting the budget if some features were of great importance. The atmosphere of trust was created while discussing the prototype, and was strengthen after implementing the core system in organizational practice.
4. Agreement on the core system was reached very early in the project.
5. The architecture of the system was designed on both the logical and technical levels.
6. The technical architecture was tightly coupled to the tools we used for developing the system, see below.
7. Our team was small, the same was true for the project budget. Even if we had desired using low-level programming, we couldn't have afforded it. High-level means employed in the project included Ruby-on-Rails (RoR) platform for developing the server side of the application, and Sencha's ExtJS toolkit for developing the client side of the application.

As follows from the analysis above, using agile methodology was justified. Moreover, it led to successful completion of the project. Note also, that the project described above *did not satisfy* the commonly used recommendations such as having *frequent face-to-face meetings* with the customers. Still it satisfied the list from Section 4 and succeeded.

6 Contribution of Our Work in Relation to Works of Others

This section is aimed at highlighting the contribution of our research while validating it through or contrasting it to the works of others. The main contribution of this paper is the *models of TSD and ASD* that highlight the difference between them and drawbacks and advantages of each, as well as the list of requirement on the P, S, and C systems that need to be fulfilled when ASD is employed. We also composed a list of practical tasks in which our results could be of use (see Section 4). We also made a test of usefulness of our suggestions in one of the practical tasks - analysis of completed projects (Section 5). This test uncovered some problems that were missed in analysis made by others.

Many of the assumption used in building the models, and conclusions derived from them are known and discussed in the literature, which we consider as an external validation for our work. For example, the weakness of the traditional development in the dynamic world is summarized in [1] as "requirements change at rates that swamp traditional methods". The reliance on tacit knowledge in agile development is well known, see, for example, analysis of literature related to the topic in [19]. The difference in usage of tacit and explicit knowledge between traditional and agile development is discussed in [13].

The fact that many of the conclusions we made based on our models can be found in various sources *cannot* be considered as *weakness* of our approach. Though an expert can identify whether a particular situation warrants a particular method based on his/her tacit knowledge, this is *not true for the whole SE community*, which is manifested in the *high percentage of SE projects failures, including failing ASD projects* (see Section 5.1).

The advantage of using our models for deriving conclusions is that while they are quite *simple, they are powerful* enough for the task at hands. In particular, they were suitable for deriving requirements on P, S and C systems that could improve chances for success of employing the agile methodology of software development. Even if we consider that all six requirements listed in Section 5 are not new, and can be found in various other sources, having them *derived from analysis of two simple models has value*, including a *pedagogical* one. It may help to convince people to actually use the requirements and not rush to using agile software development just because it is in fashion right now. An additional advantage of these models is that they are *generic and do not consider any particular brand of traditional or agile methodology*. The decision makers do not need to understand the ideas behind Waterfall, V-model, XP, Scrum, etc. to understand the diagrams in Fig. 2 and 3, and what follows from them.

Our approach to building the models of TSD and ASD is based on the knowledge transformation perspective. Though this perspective is well known in management science [8], to the best of our knowledge, it was never fully exploited in connection to SE. For example, the *only* paper we found that considered software as knowledge embedment was [20]. The paper argues that software is not a product but a new kind of storage for knowledge of which it counts five types: DNA, Brains, Hardware, Books and Software. The usefulness of our models can serve as an *inspiration to taking knowledge transformation perspective in SE research more seriously*.

Our future plans includes gathering more empirical data to support the assumption of usefulness of the ECEA and SEA models suggested in this paper, e.g. via case studies. Here, we are especially interested to see which type of tools agile teams use to enhance their productivity. Another topic worth investigation is how the knowledge obtained in one ASD project is transferred to other ASD projects considering the desire of not having any extended documentation. Our working hypothesis in this respect is that such knowledge is being built into the tools used in ASD.

References

1. Highsmith, J., Orr, K., Cockburn, A.: E-Business Application Delivery, pp. 4–17 (2000), http://www.cutter.com/freestuff/ead0002.pdf
2. Shore, J., Warden, S.: The art of agile. O'Reilly (2008)
3. Weaver, M.: Do you agree or disagree that Scrum is not Agile (May 2011), http://www.linkedin.com/groups/Do-you-agree-disagree-that-81780.S.52354777
4. Bider, I., Karapantelakis, A., Khadka, N.: Building a High-Level Process Model for Soliciting Requirements on Software Tools to Support Software Development: Experience Report. In: PoEM 2013, Riga, Latvia. CEUR, vol. 1023, pp. 70–82 (2013)
5. Agile Alliance. Manifesto for Agile Software Development (2001), http://agilemanifesto.org/
6. Conboy, K., Fitzgerald, B.: Toward a conceptual framework of agile methods: a study of agility in different disciplines. In: Proceedings of the 2004 ACM Workshop on Interdisciplinary Software Engineering Research, Newport Beach, pp. 37–44 (2004)
7. Conboy, K., Fitzgerald, B.: Toward a conceptual framework of agile methods. In: Extreme Programming and Agile Methods-XP-Agile Universe 2004, pp. 105–116 (2004)
8. Nonaka, I.: A dynamic theory of organizational knowledge creation. Organization Science 5(1), 14–37 (1994)
9. Bider, I., Johannesson, P., Perjons, E.: Design science research as movement between individual and generic situation-problem-solution spaces. In: Organizational Systems. An Interdisciplinary Discourse, pp. 35–61. Springer (2013)
10. Andersson, T., Bider, I., Svensson, R.: Aligning people to business processes experience report. SPIP 10(4), 403–413 (2005)
11. Conant, R., Ashby, R.: Every good regulator of a system must be a model of that system. Int. J. Systems Sci. 1(2), 89–97 (1970)
12. Flood, R.L., Carson, E.R.: Dealing with complexity: an introduction to the theory and application of systems science. Springer (1993)
13. Boehm, B., Turner, R.: Balancing Agility and Discipline. Addison-Wesley (2004)
14. IbisSoft, ibisSoft (2012a), http://www.ibissoft.se/node/7
15. Lacey, M.: Agile Aliance (2008a), http://conferences.agilealliance.org/sessions/20083058
16. Lacey, M.: Mitch Lacey & Associates (2010), http://www.mitchlacey.com/resources/working-software-is-not-enough
17. Lacey, M.: InfoQ (2008b), http://www.infoq.com/presentations/A-Story-of-Project-Failure-Mitch-Lacey
18. Krigsman, M.: ZDNet (2008), http://www.zdnet.com/blog/projectfailures/agileanatomy-of-a-failed-project/1100
19. Wendorff, P., Apshvalka, D.: The Knowledge Management Strategy of Agile Software Development. In: ECKM, Reading, UK, pp. 607–614 (2005)
20. Armour, P.G.: The Case for a New Business Model. Is software a product or a medium? Communications of the ACM 43(8), 19–22 (2000)

Supporting Evolving Organizations: IS Development Methodology Goals

George Bravos[1], Jānis Grabis[2], Martin Henkel[3], Lauma Jokste[2], and Jānis Kampars[2]

[1] Fresh T Limited, UK
[2] Institute of Information Technology, Riga Technical University, Kalku 1, Riga, Latvia
[3] Department of Computer and Systems Sciences, Stockholm University, Forum 100, SE-16440, Kista, Sweden
gebravos@hua.gr,
{Grabis,Lauma.Jokste,Janis.Kampars}@rtu.lv,
maartinh@dsv.su.se

Abstract. Organizations operate in increasingly dynamic and diverse environments. This leads to high variability within the organization, and the necessity to create context-aware information systems. As a part of these systems, context and business process metrics need to be monitored and acted upon. Capability Driven Development (CDD) addresses these challenges and integrates organizational development with IS (Information System) development. This paper integrates results from an industry survey and three industrial cases to define the key methodology goals for CDD.

Keywords: variability, context, capability driven development.

1 Introduction

Organizations operate in dynamic and diverse environments. They adjust their capabilities and create new business services, acquire new markets and strive to increase the number of customers. While a changed environment can create new business opportunities, those opportunities can only be reaped by organizations that have a structured approach for performing the needed adjustments. Being able to adjust to changing business environment leads to high level of variability [1]. This variability can backfire, leading to, for example, complicated product line management [2]. Moreover, there is a risk that variability is managed in low levels of abstraction when it comes to software development. This can lead to increasing development and maintenance costs. The issue at hand is thus to find means to cope with, and even thrive from, changing environments without aversively affecting the organization and its Information Systems (IS).

An organization can handle changing contexts by being prepared for changes, detecting changes, and reacting on them. Acting upon the context changes is crucial for providing adequate and usable services [3]. A part of being prepared for changes is to know the organization and its software systems in order to be able to perform re-configurations. Examples of approaches that can help organization to prepare are to

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 158–171, 2014.

employ enterprise architecture (EA) [4] and agile software development methods [5]. Having a focus on software product lines [2] or re-configurable patterns [6] is another way to be prepared for changes. Designing a set of key performance indicators [7] (KPIs) and use of technology advances such as the Internet of Things (IoT), that facilitates access to context data, enable early detection of changes.

In this paper we examine the current industry state regarding the ability to adapt to changing environments, and define a set of concrete key goals to be achieved in order to have methodological support for creating software systems that can adapt to changes. The survey focus on the organizations' use of variability management, such as the use of enterprise modeling methods and patterns. To derive detailed goals for methodological support we furthermore use three real-world industrial cases. These cases illustrate the importance of adopting to changes and allow tracing high-level methodological goals to each case.

The examination of the current state and methodological support will lay the foundation to a capability driven approach for software development. A capability driven approach supports organizations in describing the context, creating patterns to manage variability and, to deliver new services [8]. The approach, as described in this paper, is developed within the FP7 project CaaS (Capability as a Service). The overall objective of the CaaS project is to design and develop methods and tools that enable organizations to analyze, hone, and adapt their business and IS capabilities.

The reminder of the paper is organized as follows. Section 2 describes the research design, containing subsections devoted to survey design and each of the use cases. Section 3 summarizes goals from the each use case and presents survey results. Section 4 is devoted for the resulting CDD goal model and its trace to use cases and survey results. Some concluding remarks are presented in Section 5.

2 Research Design

This section describes the survey design and also covers three use cases, identifying their current IS design challenges.

2.1 Survey Design

The purpose of the survey is to provide insight into the usage of organization change management, usage of context awareness and digital services. Results reflect the actual needs of business and describe the environment in which they operate. The target group contains European companies that develop, provide or use advanced IT services. The survey is created using content management system Drupal 7, that allows localization of the questionnaire, provides comprehensive survey usage statistics and simplifies processing of results. To improve the conversion rate, the survey website is developed following the responsive web design approach and usability testing is performed prior to survey distribution. The survey was distributed via the contacts of the CaaS project to respondents in Latvia, Germany, Spain, Portugal and Sweden. The respondents were approached either as being a direct contact of one of the CaaS project partners (60 respondents), or as being a part of a professional IT related network (5 networks were approached, with a total estimated size of 200 persons). This resulted in responses from 42 organizations. The main objectives of the survey are:

- To analyze the level of Enterprise Modeling and Enterprise Architecture usage and maturity in industrial companies;
- To identify what causes the changes in companies' business processes and information systems and evaluate the degree of variation;
- To assess the level of goal and goal related KPIs usage in different companies;
- To measure the context awareness and types of context that are being used;
- To analyze the level of knowledge management artifact usage (e.g. best practices, patterns, experiences);
- To assess the level dynamic system adjustment and effectiveness of variability management;
- To identify most important features for development methodologies.

The level of *enterprise modeling* maturity is determined by asking questions about enterprise models and architecture frameworks used in the company. The respondents are asked whether they maintain their models in static or dynamic form and how well do the current models cover the whole enterprise. They are able to choose from several well-known frameworks (based on [9]).

To evaluate the business and information systems *change management*, respondents are asked to specify the frequency of changes made to their systems, type of changes, change drivers (based on [10]). Furthermore, they are asked to identify the most affected business processes (based on [11]) and their current satisfaction with implementation of software changes.

Goal and KPI usage is determined by asking questions about the business level at which goals are specified, whether or not measuring of goals is done using KPIs and the time lag of updating the KPIs. Users are also asked to specify top environment variables and KPIs they would like to put in their performance monitoring dashboard. These questions are mainly based on [12] and [7].

Context awareness is measured by including questions that aim to specify the extent to which context aware services are used and types of context employed currently and potentially. Among the proposed context types are geographical location, device type, temperature, social networking indicators and data from external sources. Users are able to specify their own context types.

To evaluate the level of *knowledge management artifact*s usage respondents are asked to name currently used reference models and reusable best practices in information system development, if any. These are open questions, thus no pre-defined response options were included.

The level of *automatic adjustment* is measured by asking to specify currently used design-time and run-time methods for adapting information systems. Among the given options for design-time adaptation were custom reporting, user interface adaptation, model based reconfiguration, custom reconfiguration settings and executable business processes. To describe the run-time adjustments being used, response options are given as follows – dynamic web service binding, adaptive workflows, intelligent systems and automatic user-interface adaption. Respondents are able to specify their own design-time and run-time methods. The variability of software products is measured by asking to describe the level of customization, involved parties and product line management.

Lastly the survey contains questions asking to specify information system development methodologies currently used in the company, and what the respondent sees as the most important features for the development methodology.

2.2 Description of Use Cases

To complement the survey with detailed samples of organizations working in changing environment three industrial cases were used. These use cases where described based on semi-structured interviews. Use cases deal with energy distribution services, e-government services and port compliance management services for ships. They illustrate different companies that operate in rapidly changing environment and seek improvement in currently used approaches for developing new services, dealing with high variability and large amounts of context data. Companies are further referred to as C-1 (energy distribution services), C-2 (e-government services) and C-3 (maritime compliance services for cargo ships).

2.2.1 Energy Distribution Services

C-1 acts as an independent software vendor and as a business process outsourcing service supplier. It has developed a proprietary software product that provides support functions especially for the energy distribution industries (e.g. grid operators, suppliers). The company operates in different markets and business context is affected by different regulations, bylaws and various circumstances, leading to increased complexity of the software product due to high level of variability. Provided services may not only be driven by external regulations but also by business opportunities that may be related to technological advances (e.g. cloud computing, IoT) and changing customer preferences. This motivates the need for a context-aware solution that is able to deliver the business value to an ever changing market.

The underlying business scenario deals with the exchange of data between two standardized energy distribution market roles – the grid operator and the energy supplier. They exchange various types of messages among which are Metered Service Consumption report messages (MSCONS) and Utilities Master Data message (ITILMD). By regulatory requirements, all data must be sent by e-mail and its format must comply with the international Electronic Data Interchange for Administration, Commerce and Transport (EDIFACT) standard and existing national variants of it. EDIFACT message formats are subject to periodical change by the regulatory authorities, with usually two releases per year. During the processing of messages different types of exceptions with varying importance might occur. Rules, that capture expert knowledge of the domain specialists, are defined to resolve recurring operation tasks, thus increasing the degree of process automation. Invalid messages must be processed by a limited number of knowledge workers. The required level of manual human interaction is minimized by using automated means whenever applicable. MSCONS messages processed by energy supplier can be checked and validated programmatically, however if any errors are found knowledge workers have to interact. Processing of ITILMD messages by the grid operator also can involve manual work (knowledge workers), however in this case it has been minimized by using fully automated rule based BPMN (Business Process Model and Notation) processes.

2.2.2 E-Government Services

C-2 maintains a service catalog that is used by 250 local municipalities with up to 200 services. During the last 5 years the overall number of platform users was more than one million. The e-service catalog contains fully automated services and ones that require some actions to be carried on at the municipality.

Services are customized for each municipality according to its needs and legislation, which is done on the level of software code. This leads to high level of variability and complex maintenance. Each municipality has its own webpage that contains a listing of available services, having a few of them promoted to the front page. The range of available services as well as the set of promoted ones differs from one municipality to another. It's important to improve the usage of e-government services and it has been observed that the decision which services to promote on each of individual websites has significant effect on it. Choosing the right services is a complex task affected by multiple factors. Judging just by the current service usage data of the corresponding municipality can achieve only limited success. Ideally the increase of demand for a specific service has to be predicted in advance, and the services need to be promoted or demoted proactively. There is a lot of information involved in achieving this – data about historical and real-time service load from the specific municipality and similar municipalities in terms of context (e.g. population, geographical location), weather and traffic data, indicators from social media channels, schedule of related events and more. Another important factor is to keep services running at all times even in case of a rapid increase in demand for a specific service. There is also a need to perform certain adjustments in services during run-time based on context – for example while booking a specific public pool online, users should be offered to book the nearest available pool in cases, when the desired pool is expected to be fully occupied.

2.2.3 Maritime Compliance Services for Cargo Ships

C-3 specializes in software and services for maritime industry and is one of the leading maritime IT companies. It has successfully completed over 1600 projects and has more than 500 customers globally. The company is also actively participating in multiple EU research projects and has secured several important distinctions and awards in maritime research. One of the services provided by C-3 is ensuring regulatory compliance for vessels, falling under the umbrella of Electronic Maritime Strategic Framework (EMSF). The EMS framework includes a number of services of great variety that C-3 aims to provide to port authorities as well as vessels. Such services will enable networking and computer supported co-operation between the principal maritime stakeholders (namely vessels, vessel owners, crew and port authorities), focusing on (i) safety and security of maritime transport services and assets, (ii) environmental protection, and (iii) sustainable waterborne transport services into efficient and secure door-to-door transport services.

C-3 mainly focuses on the tight integration and co-operation in the fragmented field of regulatory compliance in the maritime domain. Regulations are created by numerous different bodies, with little co-operation between them. As such, there is a significant lack of cohesion between the vast array of regulations and the possibility of conflicting regulations is very real. Every shipping company must comply with the regulations of a particular port, involving a large number of documents about various aspects that are of relevance to approximately 15,000 ports world-wide. Failure to comply in a given rule may have a tremendous impact in the ship owner's ability to operate. Currently compliance for Tanker Management Self Assessment (TMSA) and Planned Maintenance System (PMS) is provided, but a thorough solution to ensure overall compliance is envisioned.

3 Empirical Results

The section contains interview results from previously described use cases and analysis of survey responses. These data will be used in the proceeding section to develop a goal model for the CDD methodology.

3.1 Survey Results

The survey covered seven aspects of changing organizations: enterprise modeling, change management, goal and KPI usage, context awareness, knowledge management, automatic adjustment and development methodologies. 42 organizations participated in the survey. The organizational size ranged from small (below 50 employees) to large (over 250 employees), see Figure 1a. A majority of the respondents were active in the information technology field (Figure 1b). In the following we summarize the results of each aspect.

One of the goals for the industry survey is to analyze the level of Enterprise Modeling and Enterprise Architecture usage and maturity in industrial companies. It's important for applicability of the CDD approach for companies to be well familiar with the various principles of enterprise modeling. More than 84% of respondents have indicated that they use enterprise models. In most cases process modelling is used, followed by data modelling, organizational structure modelling and goal modelling. The survey shows that enterprise coverage by the models is average. Currently 36% percent of all respondents have indicated that only the most crucial parts are covered by models. One fourth of all responses say that majority of the enterprise is covered, while similar amount have pointed out that only the most crucial parts of the enterprise are covered by models. Although dynamic enterprise models have higher potential, static models are used almost twice as much as the dynamic ones.

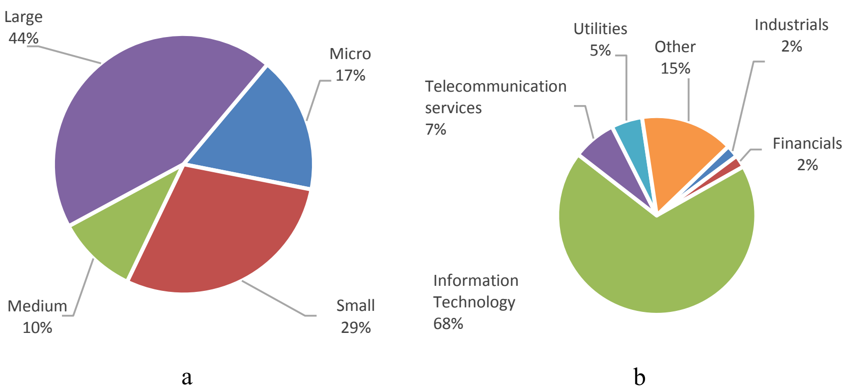


Fig. 1. Company size (a) and field of expertise (b)

Survey results show that there is a high variability of the business situation. It affects many parts of the business and most often is caused by acquisition of new markets, dealing with changing legislation regulations and having a large number of customers with unique requirements. Only 15% have responded that there haven't been any changes in the business. The most common ones are externally imposed changes (30%), followed by strategic changes (26%) and incremental changes (21%). The most affected business processes are sales and distribution (16%), production (13%), customer relationships management (11.7%), project management (9%), human resources management (9%) and knowledge processes (9%). New business opportunities (see Figure 2a) arise because of new market acquisition (33%), development of new products (28%) and introduction of new business models (26%). Lower cost (44%) and higher quality (48%) have been confirmed to be equally important efficiency improvements. Most common reasons for product requirement change are the need to acquire new markets (47%), having different requirements for different country customer/supplier market (30%) and low customer satisfaction (14%). External changes uncontrollable and unpredictable to the industry change are tighter economic conditions (31%), followed by advances in technology (26%) and new legislation (23%).

Software systems are changed for improving the existing functionality or new customer requirements (23%). Results show that new versions adhere well to the actual requirements (just 11% describe it as bad), however the situation in this area can still be improved (only 2% say that changes fully satisfy the actual need).

Identification of proper capability metrics and constant monitoring and evaluation of the capability is a central part of the CDD. This is the reason why survey contains questions about the level of use for goals and goal related KPIs in companies. The results are positive and have shown that the vast majority of the respondents use goals (85%) and measure them using KPIs (72%). Most commonly goals are specified on the strategic level (45%, see Figure 2b), followed by tactical (28%) and operation (13%) levels. The satisfaction with key performance indicator update speed shows similar trends as with software update speed in general. 35% describe the time lag as being average, 30% find it minor and 17% can't distinguish any time lag. Most common are KPIs describing key economic figures (e.g. sales margin, incomes, projects on-budget), followed by time tracking (e.g. on-time delivery, projects on time, average time to resolve an issue), quality of products and services (e.g. partner, customer satisfaction, functionality of the solution, error trend). The environment variables that most of the respondents would be willing to include in their dashboards are related to customers (e.g. customer needs, social media signals, and customer satisfaction), legislation (e.g. parties at the government, time of enforcement of a new regulation), economic factors (e.g. exchange rate fluctuations, market situation) and technology evolution.

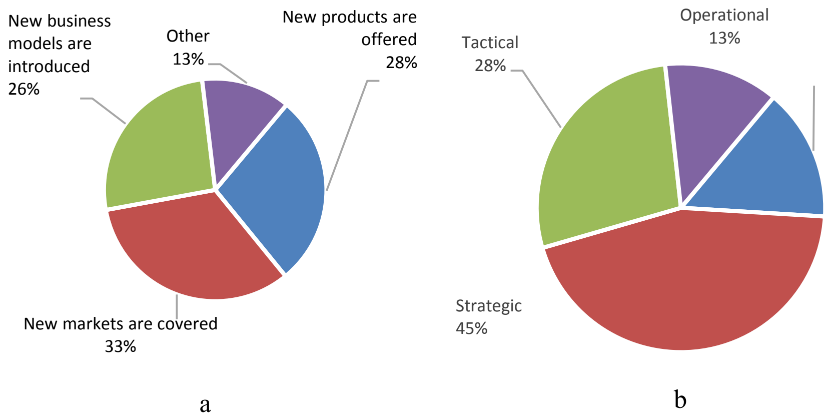


Fig. 2. The trigger for new business opportunities (a) and level of goal measurement (b)

Context data are widely used and agreed to be valuable (70% use context-aware services, see Fig. 3a). The most commonly used context-based services are location awareness and service customization for different device types. Data provided by external service providers, social networking data and sensing data are also commonly agreed to be valuable. Survey also shows the need to retrieve context-data from external service providers. The majority of our respondents (65%) have negated the use of context data to limit accessibility of information systems.

Although survey results have shown that there is a high business situation variability and software systems are being changes often, only 59% have indicated that they use methods for adapting their information systems. The most popular methods for dynamic adaptation of IS are custom reporting (29%), user interface adaptation (27%) and customer reconfigurable settings (22%). More complex methods like executable business processes, quick deployment workflows (10%) and model based reconfiguration (9%) are less popular. The survey has shown that almost 60% of the respondents don't incorporate knowledge modules in information systems. The results in the area of dynamic system adaptation are similar – 59% have responded that they use one or another method. The most popular method for dynamics information system adaptation is automatic user-interface adaptation (33%), followed by adaptive workflows (28%), dynamic web service binding (23%) and intelligent systems (15%).

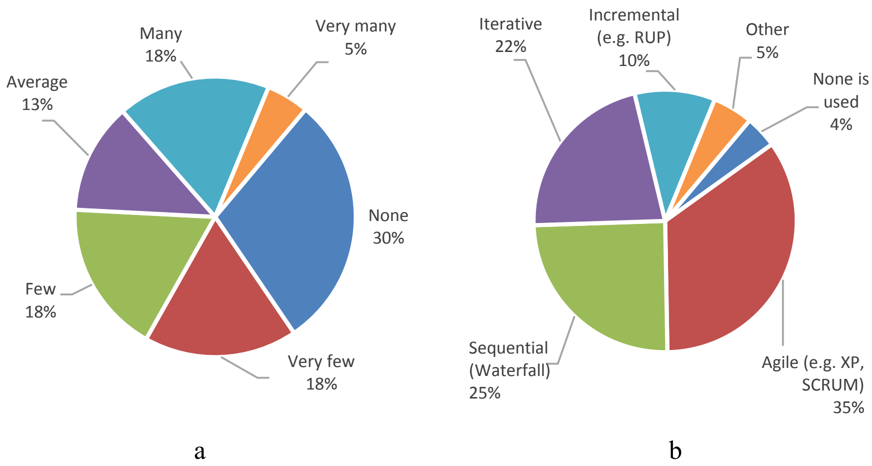


Fig. 3. Amount of context-services used (a) and development methodologies in use (b)

Only one third of the respondents arrange their products/services in product lines which conflicts with previously identified high level of variability and the need to customize products for different markets and customers. Commonly product lines are managed using product line architecture (32%), domain engineering (23%) and application engineering (23%).

The most popular formal product/information system development methodologies used by our respondents are Agile (35%, see Figure 3b), Sequential (25%) and Iterative (22%). There isn't a single most important development methodology feature as the votes are evenly distributed between all options listed in the survey.

The industry survey has confirmed that enterprise operate in dynamically changing environment. They operate in different markets, have to deal with various types of legislation regulations and have to customize products for specific customer requirements. Changes in software systems are quite common and in many cases they don't fully correspond to initial requirements. It is clear that companies are having problems while dealing with the variability and IS adjustments. Many companies don't use methods for system adaptation and even if used, the least complex ones with the least potential are chosen. The same goes with enterprise modelling – static models are more popular. Use of context data is quite common however it should be used more widely (e.g. for system adaptation). Survey shows that in most cases companies already have goals and corresponding KPIs defined.

3.2 Use Case Results

A series of interviews with C-1, C-2 and C-3 were performed to comprehend the IS development related issues they are facing daily and to collectively specify their individual goal models for the desired software system development methodology.

3.2.1 Energy Distribution Services

The main goal of the C-1 is to increase business value delivered to its customers and it is being supported by all other goals. To achieve this, proper KPIs have to be identified and constantly monitored.

The corresponding methodology should increase the level of business process automation thus minimizing costs. Currently the amount of invalid messages that require manual process is 10% - 90% depending on the customer. There is a lot of room for identifying and implementing exception handling patterns that would assist knowledge workers or even eliminate manual activities in processing of invalid messages. Manual tasks should be intelligently distributed between the knowledge workers ensuring that they are able to perform them.

Context data, like compliance with changing standards (e.g. message formats, message processing in predefined time window), must be taken into account as it's a significant factor affecting delivery of business value. The methodology should be development platform agnostic – companies have to be able to stay on the existing platform.

The shift from database-centric architecture towards service-oriented architecture should be promoted. The methodology should reduce the complexity of the product, effectively deal with variability and simplify development of vertical solutions.

3.2.2 E-government Services

The main goal of C-2 is to improve the usage of the services they provide for different municipalities. From 200 services provided only one half is in active use. The development methodology should support creating smart services (e.g. automatic update of input fields on a web-form). The smart services are services which are easy to extend and adjust to different context. They predict potential difficulties and preemptively make corrective and adaptive actions [13]. As mentioned earlier, the set of promoted services have critical effect on the overall platform usage, therefore service promotion is a task of high importance. Services have to be context-aware (both internal and external context data sources are considered), their usage data needs to be constantly monitored and potential changes should be identified in advance. Understanding the reasons for changes in service usage also allows to provide adequate IT infrastructure (e.g. automatically scale the services that are supposed to get a usage surge). The methodology should promote usage of various patterns and knowledge artifacts. The potential areas of pattern usage are service promotion and context-data based run-time adjustments, increasing the level of service automation (e.g. responses from municipality could be sent automatically in certain scenarios).

It's important to reduce the time-to market by streamlining service activation (improving interactions between the C-2 and the municipality). Services should comply with the legal regulations of the particular municipality at all times. The methodology should support effective change management process and integration with external systems.

3.2.3 Maritime Compliance Services for Cargo Ships

The main goal of C-3 is to improve the usage of currently provided compliance services and increase the enterprise's share of the market. The development methodology should promote development of smart services [13] as they provide competitive advantage.

More specifically, a framework to provide compliance services for managing regulations digitally and creating services for all the different stakeholders, will (i) enhance co-operation between the different stakeholder groups, (ii) improve the quality of regulations, and (iii) reduce the burden for those having to enforce the regulations as well as those who must comply. The services provided should also be informative regarding all current active regulations and allow preparation for regulations that will take effect in the future. Legislation issues are likely to be related to several different parameters, and services should provide for instance standard frameworks for freight information exchange, common transport documents for carriage of goods, simple and harmonized border crossings procedures.

The developed system should comply with legislation regulations from all ports. Inputs from different ports regarding legislation should be aggregated and critical legislation (regarding medical, environmental, cargo restrictions etc.) issues must be identified. Compliance with all legal issues is critical for maintaining the current clients.

The end-users should be able to easily find the service they need and most important services should be promoted. Service activation should be streamlined, improving interactions between the vessel's software system, port and vessel's officers. Services should be automated whenever possible (e.g. if a cross-check at the medical state of the crew has been carried out, it may automatically be used as an input to the system). In order to achieve this change patterns should be identified and used during run-time. The development platform should also provide effective means for integration with external systems and automated scaling of services.

4 Goal Model

The developed goal model is shown in Fig. 4 and is based on the results from the industrial survey and use cases. G-1 is related to the C-1 use case and is one of the most important goals of the particular use case. G-2 and G-16 are strongly motivated by the survey which clearly shows that companies are having problems while dealing with variability and product line management.

C-1 has identified the need to reduce complexity of roll-outs, necessity to develop vertical solutions and manage implementation of changes more efficiently. Better change management process is also mentioned in the C-2 use case. The mechanisms supporting the capability delivery adjustment are integral part of the CDD approach. This promotes the level of process automation, facilitates conformance to SLAs (Service-level agreements) and changing requirements, allows to provide smart services. The goal can be traced to all use cases and survey results, which show that the full potential of run-time adjustments isn't being realized. All use cases have shown compliance with regulatory requirements to be a serious issue, therefore G-11 has been defined. G-12 is extremely important for C-2 and any other companies acting as service providers. All of the use cases show that companies strive to minimize the amount of manual tasks and increase the level of process automation (G-13). Development of new services (G-14) is based on the goals from the C-1 use case. It is important to provide integration with both currently used internal systems (G-17) and external systems (G-15). Integration with IS of municipalities, sponsor systems and use of national services (e.g. electronic signatures) is mentioned in the C-2 use case. Integration with external systems (e.g. any other systems used by vessels or ports) and legislation data aggregation from external data sources are one of the goals from the C-3 use case. Two types of resources that can be allocated to process execution tasks (G-18) are distinguished – computational resources (G-28) and human resources (G-29). The necessity to dynamically manage these is motivated by all use cases. Dynamic resource allocation and run-time capability delivery adjustments also help to meet performance and SLA requirements (G-19). This is mainly motivated by the C-1 use case – messages, which are being exchanged between different roles of the energy distribution market, need to be processed and acted upon in a certain time window (defined by legislation regulations). The C-1 and C-2 use cases both contain goals aimed towards constant monitoring of the process and provision of analytical services, therefore G-21 and G-27 are defined. This strongly correlates with survey results, where most of the companies have acknowledged definition of goals and related KPIs. G-21 and G-26 can be traced to all use cases and the industrial survey, which also shows the popularity of context data and context-aware services. All use cases show that patterns should be identified and stored (G-22) for applying them whenever possible (G-25). This simplifies development, product maintenance and allows to provide higher levels of process automation. Goals related to predicted changes and acting proactively (G-23) are contained in all three use cases.

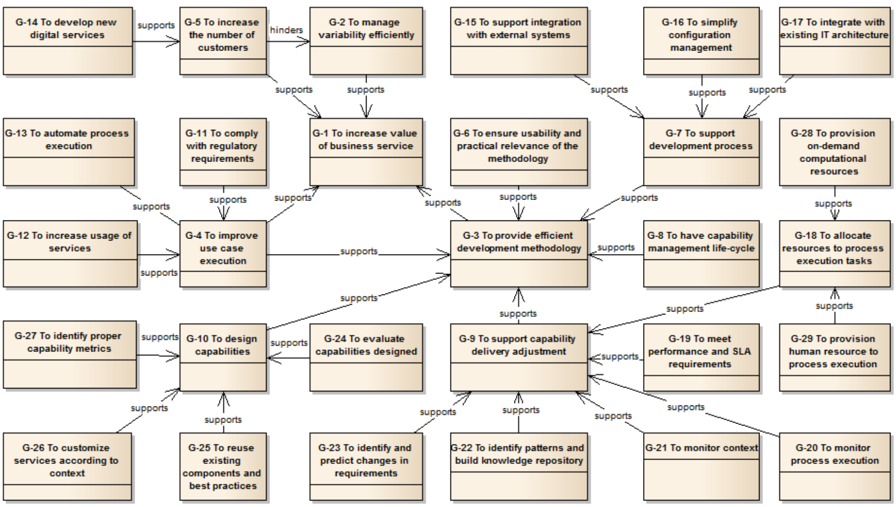


Fig. 4. CDD goal model

5 Conclusion

This paper further elaborated the CDD approach first presented in [8]. We have identified the current industry needs using survey and case studies. It has been found that companies are operating in dynamic environments and they are subject to frequent changes. Moreover, methods for dynamically adjusting the IS are not being used to its full extent. Companies are furthermore having problems while dealing with variability. Although the survey indicates that the concept of context is used widely, it could be used even more to assist in IS adaptation.

The CDD goal model that addresses the issues mentioned above was defined in this paper. There are several goals which can be distinguished as they are motivated by both survey results and use cases. One of these is G-11 - legislation changes - has proven to be one of the common drivers for software system changes, therefore it is very important to support compliance to regulatory requirements on methodological level. The empirical results show the importance of context data, motivating definition of G-21 and G-26. G-9 and its supporting goals are aimed towards providing run-time adjustments, which are also proven to be important. Process execution automation (G-13) is motivated by all use cases and should be equally important to all businesses employing process-based services.

The future directions of the research include development of the technical and methodological components of the CDD approach. Applicability of the CDD will be verified by improving the case studies contained in the paper.

Acknowledgements. This work has been partially supported by the EU-FP7 funded project no: 611351 CaaS - Capability as a Service in Digital Enterprises.

References

1. Rabiser, R., Kepler, J., Wolfinger, R., Grunbacher, P.: Three-Level Customization of Software Products Using a Product Line Approach. In: 42nd Hawaii International Conference on System Sciences, vol. 09 (2009)
2. Rocha, R.S., Fantinato, M.: The use of software product lines for business process management: A systematic literature review. Information and Software Technology 55(8), 1355–1373 (2013)
3. Bolchini, C., Schreiber, F.A., Tanca, L.: A methodology for a very small data base design. Information Systems 32(1), 61–82 (2007)
4. Boh, W.F., Yellin, D.: Using enterprise architecture standards in managing information technology. Journal of Management Information Systems 23(3), 163–207 (2006)
5. Cao, L., Ramesh, B., Abdel-Hamid, T.: Modeling dynamics in agile software development. ACM Transactions on Management Information Systems 1(1) (2010)
6. Hadaytullah, G.A., Koskimies, K.: Pattern-based dynamic maintenance of software systems. In: Proceedings - Asia-Pacific Software Engineering Conference, vol. 1, art. no. 6462707, pp. 537–546 (2012)
7. Eckerson, W.W.: Performance Dashboards: Measuring, Monitoring, and Managing Your Business. Wiley, Hoboken (2010)
8. Stirna, J., Grabis, J., Henkel, M., Zdravkovic, J.: Capability driven development - An approach to support evolving organizations. In: Sandkuhl, K., Seigerroth, U., Stirna, J. (eds.) PoEM 2012. LNBIP, vol. 134, pp. 117–131. Springer, Heidelberg (2012)

9. Urbaczewski, L., Mrdalj, S.: A comparison of Enterprise Architecture Frameworks. Issues in Information Systems 7(2), 18–23 (2006)
10. Kallio, J., Saarinen, T., Salo, S., Tinnila, M., Vepsalainen, A.P.J.: Drivers and tracers of business process changes. Journal of Strategic Information Systems 8, 125–142 (1999)
11. Mooney, J.G., Gurbaxani, V., Kraemer, K.L.: A Process Oriented Framework for Assessing the Business Value of Information Technology. The DATA BASE for Advances in Information System 27(2), 68–81 (1996)
12. Popova, V., Sharpnskykh, A.: Formal modelling of organizational goals based on performance indicators. Data and Knowledge Engineering 70(4), 335–364 (2011)
13. Allmendinger, G., Lombreglia, R.: Four strategies for the age of smart services. Harvard Business Review 83, 131–145, 158 (2005)

Consistency Issues in Large Business Process Model Environment, a Case Study

Tomas Bruckner, Vaclav Repa, and Dusan Chlapek

Department of Information Technology at Faculty of Informatics and Statistics
University of Economics, Prague, Czech Republic
{bruckner,repa,chlapek}@vse.cz

Abstract. Consistency is a quality required by many approaches to the design of systems. In our practice we have found that keeping the business models consistent is one of the most difficult and critical tasks in corporations. Therefore we decided to explore this issue and its causes in a qualitative study. This paper discusses approaches to mutual consistency between business models, and proposes original concept of weak and strong mutual model consistency, independent on modeling language. The consistency is studied in one large corporate business process model and its evolution through the years by a case study and ethnographical observation methods. Relevant model types and its important consistency relations are described. Then an influence of various factors to consistency is narrated as a result of observations. Three factors with impact to mutual consistency are drawn: complexity, relation between evolving business and stagnant model, and modeling purpose.

Keywords: Consistency, Business Process Management, Business System, Business Process Model, Business Analysis, Corporate Governance, Enterprise Architecture, Event Process Chain, Organization Graph, Aris.

1 Introduction and Motivation

The main problem we address in this paper is the fact that many factors counteract the ability to keep the process models (and other connected models) of business systems corresponding to the current status of the system, particularly the ability to keep models mutually consistent. While modeling methods and standards could be developed in hardly sophisticated manner, usage of the methods in real and complex businesses brings difficulties.

Our team applied several business process management and modeling methods in variety of organizations and we face the difficulties by ourselves. We decided to investigate the nature of the difficulties in academic way.

The problem consists of two parts. The first, theoretical part of the problem is the conceptual nature of the consistence. We should state and understand the concept of mutual consistence in general, in the relation to system modeling, and in concrete connection between selected types of models or standards. The second part of the

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 172–186, 2014.

problem is the real business system ability to keep its models consistent, the situations that can negatively influence the consistency of models, its causes and consequences.

We hardly believe that the discipline of business process management (BPM) is no less related to the management than to computer science. Even if the BPM can be used to design software systems, we believe it is necessary to be much more oriented to business systems, consisting of living people interpersonal relationships. Therefore we try to adopt a perspective of social sciences as well, in our research. We believe that methods and conclusions without field research in business could lead to confusion [1].

2 Method

We tried to find some way of quantitative research to understand the described problem, to no purpose. We have found that even if the relevance of the research could be higher, the generalization by quantity of cases would flatten the nature of the examined problem. Therefore we consider the examination of this problem by quantitative research as not suitable. So we decided to establish our approach on a single case. Our method is inspired by Yin's case study research [2] and by ethnographic methods applied in computer science [3][4][5]. We chose one corporation, where we act as consultants and observers from about 2005 and where we have access to the huge number of records of business process models and management decisions.

In the study we describe the situation as a whole to understand the context. Then we describe our observation of partial selected views. We tried to find and describe factors which occurred in evidence and which had considerable influence to consistency of the business models. The factors are described wannabe independently and objectively, but with the confidence of our point of the view and interests which disrupts the independence and objectivity. We try to empathize to stakeholders and describe the situation from their point of view. Then we find and discuss the predecessor and successor occasions, especially from the perspective of business process management and consistency, and assume them as possible causes and consequences in a narrative.

The observation is pointed to two main objects. The first is the modeling environment and process models themselves. The second target objects are stakeholders involved in process modeling and their behavior and decisions.

We tried to take the role of an independent observer, even confident of our influence as methodic consultants in the corporation. We fully understand that the results of our observations are applicable to this and only examined case, and the causes and consequences are our interpretation. But we believe that the fact the described situations occurred in this one case is valuable knowledge that similar situations and influences just could occur. Moreover, this belief is indirectly supported by our similar experience in other companies. On that basis it is not possible to state general principles, so we generalize at least few principles to learn negative consistency factors while managing business processes.

To understand the influence of the examined factor to model consistency, we should first define our understanding of the consistency as a concept and as a problem, which is done in chapter The Consistency. Our approach to the study and the study itself is described in detail later in chapter The Study.

3 The Consistency

In this chapter we define our understanding of the concept of consistency. There is a lot of work done in the field of mutual consistency between models, within particular modeling method. The need for caring of the mutual consistency of different system models can be found almost in all traditional system development methods based on multiple views on the system like Yourdon method [6], its predecessor SADT [7], and SSADM [8][9] for instance. All they took care of mutual consistency between two different views on the system, the functional (activity, process) and data view. Lately concepts of consistency in database and transaction theory [10], or consistency within object modeling approaches like OMT [11], UML [12] or others were done. In the field of process models and their consistency with other models we feel a lag in research. Basic principles are described mostly in implementation tools guides of modeling tools or CASE tools (like Aris, PowerDesigner or Enterprise Architect), but research work is rare [13], [14].

We try to accept mentioned approaches but define the mutual model consistency independently on modeling method to facilitate using soft methods of consistency research and application of the concept to mutual consistency between completely different modeling approaches. In this paper we do not address a structural consistency within one model as it is defined in [15].

In our approach to business modeling we define inconsistency, weak consistency, and strong consistency, between two or more models of one system, as follows:

- Models are inconsistent, if interpretation of at least one model is contrary to interpretation to any other model.
- Models are weakly consistent, if interpretation of each model is not contrary to interpretation to every other model.
- If some model or set of models are intended to describe the system from some point of view exhaustively, then models are strongly consistent from this point of view only if interpretation of each other model from this point of view could be interpreted as constituent part of the exhaustive description.

This definition is based on model interpretation and on contradiction. Interpretation and contradiction are dependent on the modeling language. Therefore inconsistency could be turned to consistency by changing the meaning in the modeling language. The definition is meant to be general, applicable also to other modeling approaches than system models of even BPM.

Correspondence and Coherrence. When modeling systems, in particular in system analysis, we require certain truthfulness of the model, which means that the model corresponds with the system which describes. As we know, the system is just a way to understand part of our complicated world [16]. The consistency between models, in opposite, doesn't mean correspondence, but a coherence. Finding inconsistence, by definition, means finding contradictions between models. If we assume the complicated world is coherent, or, at least, if we require the examined system should be coherent, we must claim that the models of the system should be coherent too. It means that when we require consistency between models, we require different models to describe the same system. By creating different models we try to raise one system from the complex world by more different views. That is why the consistency is important. If we would not check the consistency, probably every model (or modeler) could describe different system from the others which would lead to misunderstanding at least.

Weak Consistency. The weak consistency means that different models are not in mutual contradiction. If we assume two or more models are intended as models of one system, they are weakly consistent if we interpret each model the way which is not in conflict with interpretation of any other model.

Strong Consistency. The concept of strong consistency is more difficult. We must assume, that we model the system from different points of view, probably by different languages, and that there is an intention to describe the system from one of the points of view exhaustively, sometimes we say globally. Then, if we interpret every other model from the same point of view by a way, which we can reinterpret as a component of the mentioned global description, the models are strongly consistent. (Let's see an example, if we keep an organization chart as an exhaustive model of the whole company organization structure, and we have process model, where is a worker connected as an actor to an activity, then we can consider the process model and the organization chart as strongly consistent from the organization point of view only if we are able to interpret the worker as part of the organization structure.) We can see two interpretations there, the interpretation of a model from a certain point of view, and second, a reinterpretation of the first interpretation to the global model. This exempts the definition from the addiction to one common modeling language and to usage of the same elements or connections in booth models. (Then the process model and organization chart can be modeled by different modeling approaches and the actor can be once modeled as a person or a position and second as an organization unit, which it is a part of. It is only necessary to choose the point of view and to be able to interpret).

By the term *model* we understand in general such representation of a certain thing, which has some similar attributes as the thing, so that it is possible to get some knowledge about the thing by the representation [17]. In the field of business system modeling, by the term model we understand usually a diagram or a set of diagrams which in some modeling language describe the system, at least the elements of the system and their relationships.

4 The Study

In this chapter we describe our findings gathered by the long term observations and analyses in the real company. Firstly, the overall situation is described, five relevant types of connections in the system are enumerated and the way which they influence the system's consistency is explained. Then main observed factors with negative influence to the consistency of the system are summarized, analyzed and finally generalized.

4.1 Overall Situation

The company where we performed the study is an international corporation. We don't have permission to publish its name or identification. We investigated few companies included in the corporation, situated in central European countries.

Number of the companies of the corporation varied because the corporation structure was redesigned due to European legislation [18][19] and shareholder structure changes. The legislation requested to unbundle company parts concerning the regulated business out from the parts concerning unregulated customer sales, to free up the market and enable third party traders to use the regulated parts of the business as a service. Furthermore, the corporation absorbed several regional companies. This was an occasion first to merge the purchased companies and second to cut it into regulated and unregulated parts. A similar situation was performed inside the corporation in different European countries, which are as much as possible independently managed units with some shared services. We describe the observation in one country.

From the business process management view, it was necessary to find the common activities and standardize processes, and then find suitable borderline between unregulated and regulated business, cut the processes and define the border as a service. This situation brought an opportunity and a need to create consistent business process model to make the unbundling decision precious and reasonable.

We acted in the described situation as consultants in methodology of business change documentation from 2005 till 2013 with varying level of engagement. We are fully aware that described observation is determined by this our role, therefore it is not unbiased, it is our interpretation.

4.2 The BPM Situation

Most of the companies in the corporation have established a unit responsible for business processes and organization (P&O). One of them keeps a methodical function, and provides methods to analyze and document processes and other parts of corporate architecture (like directives, guidelines, organization, services, software applications etc.) to other companies. The P&O department is included in one unit together with project office and ICT. P&O is meant as a service for top management to keep documented business structure and to track organization changes, and as a service to employees, middle and low level managers to see consequences of their work and find and communicate partners of intended business changes within organization.

While the company structure is hierarchical, the corporation tries to analyze, document and manage end to end processes passing through more organization units and sometimes even through more companies.

They understand that process structure should be flat instead of hierarchical. There is one level of detail selected as process level. Those processes are grouped to process groups, and can be broken down to parts of a process or sub-processes.

Through time the overall number of process models is in thousands and the number objects in the models are in hundreds of thousands. Overall number of affected employees is in tens of thousands.

As a primary modeling tool for business models documented by P&O the Aris Business Architect (formerly Aris Toolset) and Aris Business Server have been chosen. For internal publication of processes the Aris Business Publisher is used. Modeling methodology is documented in a form of Aris filter and in forms of powerpoint snapshots with modeling guidelines. For quick reference there are so-called QuickCards with common modeling techniques available.

The models are used as obligatory inputs for modeling and analysis in special cases, as for software development or information systems design, compliance audits, technical quality management etc. If these initiatives use special tools to modeling processes, their outputs are subsequently implemented back to Aris model environment.

There are process owners and process experts designated for every process. Process owner is a role responsible for management and up-to-date documentation of the process. Process owners are usually managers at higher level of organization structure, who can depute the responsibility to some of their subordinates. Process expert is a role of a person who has long experience of process performance and is able to technically consult the process definition as well as possible impact of proposed process changes.

4.3 Observed Consistency Constituting Connections

While examining the governance of mutual model consistency in given situation, we follow our definition of consistence described above. We are looking for most important connections between models related to significant aspects of our approach to consistence. The two essential ones are the contradiction between models, and the completeness of models from different views.

The connections we observed were mostly relevant to the Aris method determined by the modeling environment used:

1. Logical and time consequences between different process flow models (modeled by EPC),
2. Hierarchical connections between process flow (EPC), overall model of one process (modeled by FAD), models of value added chain (VAC) and overall model of processes (process house modeled by VAC),

3. Information flows between process models (EPC, FAD) and other relevant models (inputs and outputs of activities),
4. Connections related to organization, such as relations between workers, roles, positions, organization units, teams and external subjects, and their responsibilities of activities and processes, services and others,
5. Connections between approved models and models in change.

Ad 1) Logical and Time Consequences. The process flows are modeled in EPC diagrams of several types. Connections between two processes are in accordance to Aris method modeled by an event object and an interface symbol of function object. The event has two occurrences. The first one represents particular end state in the first model, and the second one represents the starting event in the second model. In the first model the state is connected by “is successor” connection to the end of the process flow and by “is predecessor” to the interface. In second model the event is successor of the interface and is predecessor of the first activity. The interface object represents the process, which is a function object, from which the EPC model of the process is hierarchized. We can consider the models as mutually inconsistent, if the event is not a shared definition of one object, if the event has less or more than two occurrences, if these occurrences are not in correct models or are not connected properly, if the interface is not an occurrence of the correct process or if the process is not hierarchized to correct model. If process flow on one side of the interface is not modeled, i.e. the process object is not hierarchized to EPC model, the event occurs only once.

Ad 2) Hierarchical Process Connections. As mentioned above, the processes are modeled in flat single level structure. The processes are grouped to groups and these groups to groups at higher level, which makes hierarchical structure. On the other side, one process can be described by more process flow models. Processes, Activities within process flows and groups of processes are in the modeling environment represented as function objects, which are hierarchized to relevant models as follows. The hierarchy between process groups is modeled in Value Added Chain diagrams (VAC), one model for one level of hierarchization. The process houses, i.e. maps of processes relevant to one company or factually related, are modeled in VACs or other types of models, in some cases without hierarchy, in some cases with multi-level hierarchy. On the process level, there are Function Allocation Diagrams with allocated participating companies (as organization units objects), process owners and process experts from the companies (as persons or job positions objects), goals of the process and its metrics (goal and KPI objects), important general inputs and outputs (documents, documented knowledge, cluster objects) and products or services produced by the process (product/service objects). The process can be described in process flow (EPC models) or, if the process flow would be too complicated, by VAC, where parts of the process are hierarchized to more EPCs. When process flow is not necessary, e.g. if it is up to invention of the actors how the process will be performed, the EPC is missing or replaced by FAD with relevant information other than the process flow. Function object representing process or group of processes can have an occurrence in several models. Function object representing an activity in process flow can have only single occurrence. If there are similar activities in different EPCs or in the same EPC, different object definitions should be used for every one of them.

Ad 3) Information Flows. In some processes, information inputs and outputs of every activity of the process flow are modeled. In some cases they are modeled as inputs and outputs of process as a whole in FAD. Input means information, which is necessary for performing the activity (or process) and output is information, created by the activity (or process). Inputs and outputs can be Document, Cluster or Documented Knowledge objects. The Document object is used for any information, structured or not, in any form. The Cluster object is used for information stored in a database, usually accessed by an application, which is also connected to the activity in EPC. Documented Knowledge objects are used mostly as inputs and represent directives or guidelines necessary while performing the activity. In few cases, where the guideline or directive is created, they are modeled as outputs. If there is an information flow between two processes, occurrence of one Document / Cluster / Documented Knowledge object should be used as output of activity in one process and as an input in second process. One object can be used in many occurrences as input and also as an output. The possibility of multiple output usage is enabled by generalized names of the information. Even if activity adds or changes single attribute into a complex information (e.g. item price into an invoice), the general name of the output is modeled (just the invoice). Document objects, Cluster objects and Documented Knowledge objects all have their own folders for object placement (like other objects shared between models; objects relevant to models like Function objects, Events, logical operators etc. are placed in the same folders as the models). Structure of documented knowledge is modeled in special models. Structure of applications is documented in IT Architecture model. Clusters and Documents have no special structure models in Aris environment, they are modeled specifically for a reason, e.g. when developing a software applications.

Ad 4) Organization Related Connections. Every company of the corporation had its own hierarchical organization graph (organization chart) in several forms. First the organizational structure was modeled in single level organization charts and the whole company structure was covered by a number of models. This was suitable for automated elaboration, but very unfamiliar for maintenance. So currently organization charts are multi-level and cover the whole structure of the company. Organization charts can be clear, assembled from organizational units and its subordination only. More detailed models contain organization units and their job positions or types of job positions. The positions form the unit, one of them leads the unit. Even more detailed organization charts include moreover concrete workers employed in the job positions. Teams and external subjects are modeled in special models. Other type of organization chart models an assignment of roles to job positions. Some roles are specific to a group of activities, and some are general (like employee or driver). Name of the general role starts by underscore. In process flows, as actors of any activity can be modeled objects of a role, a job position, organization unit, team or an external subject. Actors of fully automated activities are applications. Organizational objects are connected to activities by several types of connections. Over time the types of connections settled at standard RACI. Even if different types of connections (RACI) are used in process flow, when managing processes as a whole instead of activities, all organizational connections in EPC or FAD are interpreted as C, process expert connections as A and process owner connections as R. The connections between process, its activities, roles, its job positions and its organization units are used to generate RACI matrix processes/organization units.

Ad 5) Approved / In Change. Business process management needs not only current version of process models. In our case there was a need for managing complex process changes, so the processes and their descriptions were most of the time in change. In time of the biggest changes there was a concept of releases. There were planned important dates of process model releases derived from the dates of corporation structure changes, few times a year. The complete process model environment was analyzed and designed as a valid and obligatory process structure due the date of the release. Today, when the organization is more stabilized, the process model is released continuously. When there is a request for process model change, relevant models and objects are copied from Approved folder to folder In Change to elaboration. The current process in Approved remains valid, until elaborated change request is approved and the changed process integrated to the Approved folder.

4.4 Observed Factors with Negative Influence to Consistency

In this chapter we describe the factors which negatively influence the process model consistency, i.e. the important consistency connections, which we observed in our case. First we describe the observed factors as a narrative, and then (in next chapter) we try to generalize.

Many involved person work on process models. The P&O department covers the governance over the model in a company. In the period of huge transformation, the whole model was managed from one place in the corporation. Now there is an effort to keep at least methodology governance common in the corporation.

In our early observation, the process model was intended particularly to get knowledge of the corporate activities from individual companies and individual persons, to have sufficient information to decide on the future corporation structure and boundaries of new established companies. This was factor with strong influence to model consistency pushed by corporate owners, top management and a variety of players of future influence. Only several conservative influential persons or departments, which had interest to keep their knowledge in private, acted against the modeling.

In late years of our observation, as the corporation had been redesigned and new structure of corporation became stabilized, the intended purpose slowly disappeared. Now, the purpose of the model is not uniformly stated, several stakeholders use it for several reasons. We have witnessed the rise of highly consistent extensive process model, and then we felt its stagnation and slow fall. During the time we have experienced many factors influenced its consistency.

Visio and Swimlanes. Firstly a famous international consulting company was hired to analyze and model the corporation processes as a basis for anticipated restructuring. The methodology of the process description was left up to the consulting supplier, who was very pragmatic in the analysis. The analytics walked around the company, interviewed the employees in detail and documented the process flows in MS Visio BPMN process maps [20]. These maps were the output and a result of comprehensive analytic project.

The partial process maps well documented the status of the corporation and seemed very useful, unfortunately only till the moment, when the managers tried to merge and cut processes through different Visio process maps. The maps used together were inapplicable. In detail we observed that different interviewed persons see the same or following processes considerably different and the managers were not able to connect models together, even if everybody knew the modeled situation on daily basis the processes work. Other difficulty was the usage of swim lanes [20] to describe process actors. Even if every modeler would use the same structure of the swim lanes, which was not true, every change of responsibility in activities (and while turbulent restructuring the planned responsibilities changed every day several times) meant redrawing the process flow. This experience led to reject MS Visio as modeling environment and to reject the concept of swim lanes to model actors.

Sophisticated Model Environment. One of companies in the corporation documented especially technical processes in Aris environment in past, so they suggested to remodel the process maps to Aris and then solve the process connections in this environment. At that time also new P&O manager to the leading company in corporation was hired and he decided to develop and manage a strong methodology of business process model. We (the authors) were asked to remodel the initial process maps to Aris EPCs, which we was not able to do without the methodology, so we initiated and participated in its creation.

After the models were translated into EPCs, possible logical connections between processes were identified by the modelers and new interviews with involved process performers on both sides of process interface was arranged to negotiate the best possible interface documentation. These interviews had great side effect to train involved persons in developed modeling methodology and reasons of consistency between process flows.

Instead of swim lanes, in EPC the actors were documented by simple connection of the actor to the activity, as mentioned above.

Every change in process flow in EPC then was accompanied by searching relevant interfaces to other EPC model. P&O provided trained modeler to every team involved in organizational change, who kept in mind possible consistency issues and coached meeting and talks with other teams or process owners possibly affected by a change proposed by the team. The teams then were able to see huge impacts of as they believed small changes in process and the mechanism of corporation redesign was set. The mechanism ensured, that top managers could decide about the company borders on top level, and the detail work was done without their knowledge by the cooperation of individual teams, and only interface problems that failed to resolve bubbled up to the top managers.

Consistency Checklists. Accuracy and consistency checklist were created and before any change was approved, the checklist report was created and communicated with the team responsible for the change. In some cases the process team was not very cooperative and was not willing to find correct consistency problems. It was an issue with less communicating teams, which sometimes after running out of project allocated time produced models with too many consistency issues, that resolving them would take more time than complete process redesign. These situations sometimes ended up just with incorporation of the bad processes with documented consistency errors by the checklists, and sometimes by deleting the processes and statement, that process flow is not necessary in that case.

Groups of Processes. The connections between processes on top level were mapped by value added chains and processes were grouped to process groups. At the level of VAC new companies responsible for the processes were documented. A secret cockpit room with actual value added chains with responsibilities was created, where top managers could walk between the models and talk and think about the restructuring without a need of PC.

The structure of the groups of processes addressed the completeness of process model. The process structure should be exhaustive, as every activity in corporation should be subsumed to some of the process groups. The concept of the interface addressed the un-contradiction between models.

Organisation Structure and Processes. Theoretical approach would expect to change organizational structure in relation to the process design. Many of managers leading the restructuring didn't think about processes, but about people and organization units. They designed the new corporation structure by cutting organization units and reallocating people. Task of the design teams then was to find activities done by concerned units and people and try to redesign processes to match the presumptions made by the organization structure. Many managers were used to think of responsibility assignment by responsibility matrix, which take into relation processes and organization units, e.g. RACI. But his matrix was either too general either too large (special scripts to generate matrix from Aris to excel must be developed since the maximum number of rows was exceeded about six times).

Either way there was a need to model and keep consistent organization structure too. Several managers and teams had several needs of modeling organization structure. The methodology changed many times and oscillated at ability to model process responsibility by all possibilities, by role, by organization unit, by general job position, by concrete job position, by worker, by team (temporary organizational unit), and by external subject (customer, suppliers etc.). This variety probably would not be possible to model by pools and swim lanes. So methodology of organization chart structure was developed to keep the organizational structure consistent. In that complexity it was not possible to keep the consistency by people's minds, and automated consistency checks was necessary. There were developed scripts that tracked multiple connection process-activity-role-general position-concrete position-org. unit-company, with possible missing parts as activity can be connected to any actor, and creating consistency check reports and RACI-similar matrix. The consistency check reports generated results like: "process a in VAC b is connected to org. unit c, d and e, of company f and g, but in EPC h was by role i through positions j, k and l traced only org. units m and n of company o". Similar sentences were understandable for modeling teams and led to consistency improvement by correcting the right models in the chain.

In some companies and some, mostly technical, processes, the description of process flows, roles and organization was sufficient for automated generation of detailed textual descriptions of the job positions to the employment contract appendix.

Opposite to P&O department who managed the business model in Aris, there was an influent HR department, who kept alternative organizational structure in SAP HR system. The structure in SAP HR consisted only of job positions and their hierarchical connections (subordination). No organizational units were used. A decision was made, that primary register for the job positions will be the SAP HR. Then an interface between SAP and Aris was developed to keep the records consistent. On the side of the Aris then it was necessary to reconstruct organization units and general job positions from the structure and job position titles from the SAP after every update. This task has failed to fully automate so far.

Sophisticated modeling environment became a pattern to roll out the experience of BPM to other European countries in the corporation.

New Staff. The complex system of connections was very effective and impressive as long as it was highly up to date. When used to manage the change, the system was great. After the deep and fast restructuring started to finish, the leading persons of the model, P&O manager and his superior left the corporation. New managers were hired from outside of the corporation, completely without knowledge of the evolution of the model. They did not experience problems, which arose and was solved during the time, and therefore they didn't feel the necessity of several precautions. They did not preach the model consistency to managers. On the contrary, as they were new, they transposed views of influent players in the organization, which were at most one-sided, not complex. The situation resulted in decisions, which strongly attacked the consistency sustainability.

One of the decisions was to leave the strict modeling methodology. The way several teams and process owners modeled their processes was no more managed so strictly. The motivation was cost savings of internal process consulting services. The overall business process management costs were excessive in the view of new management, and after restriction of supplied internal consulting services the BPM now consumed considerably lower budget. But the mutual model consistency disappears without continuous consulting and performing checklists. After time, small process inconsistencies are designed, the resolution is not entered into model and significant part of the model must be considered as inconsistent. Process groups, which are not consistent from personal initiative of process owners, are completely deleted, as maintenance of the model become annoying and ineffective. This means the completeness and the strong consistency management become impossible.

Decay. Then, one of the decisions made, was to stop using roles to describe actors in processes. Reading process models itself and managing consistency of above mentioned chain between activities and organization units was not intuitive and possible without support of Aris automation. From manager's perspective were more suitable tools MS Excel and MS PowerPoint, and thinking about roles seemed too complicated. From the P&O and some process owner's perspective modeling of roles was unnecessary. The decision to discontinue role maintenance had significant influence to companies, who used description of roles ant their connection to processes to automated job description generation. These companies started to maintain the process model at their own, without respect to methodology governance of the corporation, with preservation of the strong process consistency, but only in the given company without connections to other corporate processes.

Next factor was company ownership influence in the corporation redesign. When final planned corporation structure meant to be approved by general meetings of shareholders of the redesigned companies, in some companies the minority shareholders refused the approval of the new structure. This had two consequences in the model consistency. First, the negotiated, planed and consistent release of corporate process and organization structure was partly valid (in companies which approved the new structure), and partly invalid (in companies which did not). The previous release of model remained in use in the unapproved part of the corporation. Of course, the old release and the new release were not tuned and partially were in contradiction, therefore inconsistent. This led to several nonstandard alternative solutions to make the release compliant. Next consequence was the intention not to manage the process structure through the whole corporation. As the redesign was intended to provide access of third parties to partial services of the corporation, the overall process management would disadvantage those third parties as they would have no such possibility to participate to end to end process optimization and management. Then the processes were cut to end up at the borders of the particular companies and management and modeling of the processes was delegated to the single companies. The companies now can use only methodical services of the original corporation P&O, which now have only minor influence.

4.5 Observed Factors Generalization

In this chapter we try to identify general reasons of the factors observed and described in the narrative above. We rethink and try to reinterpret the observed situation in most mentioned factors. We tried to generalize to factors which, with confidence of scientifically irrelevant generalization, might be considered as factors that could have negative influence to consistency in complex business process models in general. We found that much of the narrated may be generalized in following three factors:

1. Business system is too complex for a human to keep in mind all consequences,
2. Business system is evolving by its nature, while the model stays rigid by nature,
3. Absence of strong modeling purpose or interest.

Ad 1) Complexity. The main factor we experienced was that the large business system is very complex and even after comprehensive analysis it is probably impossible for human to oversee all relations. That means that every decision necessarily disregards some consequences. Every change is then a negative factor to consistency, as every change must be checked for possible consistency connections that were overseen by humans. In management view it means that central decisions without distributed multi-human responsibility over the business have negative consistency influence.

Ad 2) Evolving System vs. Stagnant Model. The business system is a social system which evolve naturally by almost by relationship of people as elements of the system (in Boulding's eighth level system classification [21]). In contrast the model, even if highly managed and automated, is in principle stagnate and its nature is to stay as it is before it is changed from outside (Boulding's level two or three). The business system necessarily runs away from the model and the model must be managed to correspond the business. In management view this factor means an absence of shared sophisticated modeling methodology and of intensive communication.

Ad 3) Modeling Purpose. The model can be kept consistent only if there is a strong will to manage the consistency. The consistency is not an attribute which would arise in itself. If the will, the purpose or the interest disappear or fade, the previously kept model consistency start to disrupt, which could significantly degrade the manageability of the model.

5 Conclusion and Future Work

We experienced and observed a unique situation in extensive business process model environment with negative factors influencing mutual consistency between models. We defined the mutual consistency by interpretation, non-contradiction and completeness, described the factors we observed by a narrative and tried to interpret it in general as Complexity, Evolving Business System vs. Stagnant Model and a Modeling Purpose.

With further research we address particularly principles and methods of business modeling governance in corporations and modeling methods adjustments to avoid negative influence of observed factors, as well as elaboration of our theoretical approach to consistency to several modeling methods.

References

1. Sayer, K., Harvey, L.: Empowerment in Business Process Reengineering: An Ethnographic Study of Implementation Discourses. In: Proceedings of the Eighteenth International Conference on Information Systems, pp. 427–440. Association for Information Systems, Atlanta (1997)
2. Yin, R.K.: Case study research: design and methods (2014)
3. Myers, M.: Investigating Information Systems with Ethnographic Research. Commun. AIS 2 (1999)
4. Sharp, H.: deSouza, C., Dittrich, Y.: Using Ethnographic Methods in Software Engineering Research. In: Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering, vol. 2, pp. 491–492. ACM, New York (2010)
5. Passos, C., Cruzes, D.S., Dybå, T., Mendonça, M.: Challenges of Applying Ethnography to Study Software Practices. In: Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, pp. 9–18. ACM, New York (2012)
6. Yourdon, E.: Modern Structured Analysis. Prentice-Hall Inc., Englewood Cliffs (1989)
7. Marca, D.A., McGowan, C.L.: SADT: structured analysis and design technique. McGraw-Hill Book Co., Inc., New York (1988)
8. Downs, E., Clare, P., Coe, I.: SSADM, Applications and Context. Prentice-Hall, Inc., Upper Saddle River (1987)
9. Ashworth, C.M.: Structured systems analysis and design method (SSADM). Inf. Softw. Technol. 30, 153–163 (1988)
10. Gray, J., Reuter, A.: Transaction Processing: Concepts and Techniques. Elsevier Science (1992)
11. Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F., Lorensen, W.: Object-Oriented Modeling and Design. Prentice-Hall Inc., Englewood Cliffs (1991)
12. Egyed, A.: Instant Consistency Checking for the UML. In: Proceedings of the 28th International Conference on Software Engineering, pp. 381–390. ACM, New York (2006)
13. Ryndina, K., Küster, J.M., Gall, H.C.: Consistency of Business Process Models and Object Life Cycles. In: Kühne, T. (ed.) MoDELS 2006. LNCS, vol. 4364, pp. 80–90. Springer, Heidelberg (2007)
14. Repa, V.: Consistency in the Business System Model. In: Druzovec, T., Jaakkola, H., Kiyoki, Y., Tokuda, T., Yoshida, N. (eds.) Information Modelling and Knowledge Bases XXI, pp. 247–257. IOS Press, Amsterdam (2010)
15. Repa, V.: Information Modeling of Organizations. Bruckner Publishing, Zivonin (2012)
16. von Bertalanffy, L.: An Outline of General System Theory. Br. J. Philos. Sci. 1, 134–165 (1950)
17. Bruckner, T., et al.: Tvorba informacnich systemu: principy, metodiky, architektury. Grada Publishing, Praha (2012)
18. Directive 2003/55/EC of the European Parliament and of the Council of 26 June 2003
19. Directive 2009/73/EC of the European Parliament and of the Council of 13 July 2009
20. Business Process Model and Notation (BPMN) Version 2.0, `http://www.omg.org/spec/BPMN/2.0/`
21. Boulding, K.: General Systems Theory. Management Science 2(3), 197–208 (1957)

Supporting Introduction of Social Interaction in Business Processes

Roberts Cerenkovs and Marite Kirikova

Institute of Applied Computer Systems, Riga Technical University,
1 Kalku, Riga, LV-1658, Latvia
{roberts.cerenkovs,marite.kirikova}@rtu.lv

Abstract. Organizations that implement BPM are constantly trying to improve and effectively manage their business processes. Taking into consideration that every business process involves a reasonable amount of social interactions, i.e., simple or complex collaboration between people that occur internally (within organization boundaries) and externally (as suggestions, feedbacks to consumed services or products, and other interactions); it is necessary to provide means for incorporation of social interaction into business processes. While the need of incorporation of social interaction in business processes is well recognized, there is a lack of simple supporting guidelines for supplementing traditional business process development approaches with incorporation of social software. Iterative and incremental social software integration approach proposed in this paper provides a set of methods and concepts that are organized across well known iterative incremental development approach. The proposed approach can be used as guidelines in social software incorporation into business processes for introduction of social interaction.

Keywords: business process, social software, social BPM.

1 Introduction

Business process management (BPM) and social software have received much attention in the academia and industry due to profound research potential and gained business benefits. Social software integration describes the homogenization of unstructured work provided by social software with the process-oriented activities essential in business process management.

BPM is a process-centric discipline which founds its roots in business process re-engineering methodology [1]. BPM provides a cross-functional collaboration. Hence, BPM is considered as a holistic management approach [2], [3]. The Association's for Information and Image Management (AIIM) survey results show that more than a half of enterprises benefit from implementing the BPM solutions [4]. According to Harmon & Wolf's reported survey every fifth of the surveyed organizations spent from 1 to 5 million in business process analysis, process management, monitoring, redesign, and improvement activities [5].

Organizations that implement BPM are constantly trying to improve and effectively manage their business processes. Every business process lifecycle step

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 187–201, 2014.

produces a reasonable amount of social interactions, i.e., simple or complex collaboration between people. These interactions occur internally (within organization boundaries) and externally (as suggestions, feedbacks to consumed services or products, and other interactions). it is common, that social interaction is more intensive in less automated processes. However, a business process that is less automated usually is also more unstructured and barely repeatable, and dependent on actor tacit knowledge. Hence, the evaluation and performance of less automated processes becomes a sophisticated task that has to take into consideration not only specific performance aspects of a process, but also human (actor) qualities such as competence, knowledge, collaboration skills, etc..

Social software is one of the ways how to support social interaction. Currently there is a lack of information available of how to systematically organize and lead towards successful social software integration within BPM. Therefore, the research work discussed in this paper is aimed at providing approach that gives an opportunity to incorporate software supported social interaction into business processes in the structured and guided way. The approach is developed by extending well known iterative and incremental development approach with the set of methods and tools for incorporation of social software into business processes.

The paper is structured as follows: Section 1 briefly discusses problems regarding the incorporation of social interaction in business processes and introduces the proposed approach. Further Section 2 to Section 6 describe the phases of the proposed Iterative and Incremental Social Software Integration (IISSI) approach. Section 7 provides the application of IISSI approach within the foreign studies planning business process. Section 8 amalgamates the contribution and provides brief conclusions.

2 BPM Social Interaction Problems and IISSI Approach

A number of issues are closely related to difficulties of social interaction handling within BPM. Some of the problems (lack of information fusion, information pass-on threshold) extend so called Model-Reality Divide phenomenon (deviation between the abstract business and the actual executed business processes), other (Lost innovation) relates to an ineffective or absence of the knowledge management activities [6], [7]. The Influence of New Communication paradigm to BPM is also discussed [8].

Lack of information fusion. The importance of involving all possible stakeholders into activities is known and has been the case for discussion continuously [9]. However, business process modeling process often lacks proper involvement from all important stakeholders, especially the business process performers. If there are no bottom-up, peer-to-peer communication channels established, performers are forced to adopt a designed process model delegated from the top management. Not every designed step is precisely followed, though. As the result, Model-Reality Divide phenomenon can be observed.

Lack of participation. It is believed that not all stakeholders are participating in the business process design and execution phases [6]. As the result, the time to plan and develop highly complex business process is increasing and hence it does conflict with the current requirements of agile enterprises.

Information pass-on threshold. The overly rigid controls on the information pass in the organization can stop users from sharing it. Moreover, excessive formalism of sent material (such as memos or letters) often creates unnecessary effort for the user. Furthermore, information processing and the decision making can also be not transparent enough to the stakeholders. These factors create an impression that message success is improbable. As users cannot share their ideas easily, potentially useful information might be lost and opportunities for business process improvement not realized.

Lost innovation. The process of identification and utilization of weak ties in the organization can have a significant influence on organizational agility and finding useful knowledge in the enterprise [10]. The existence of such knowledge, critical for possible improvements of business processes, is often unknown to the process owners. As a consequence, possible innovative ideas remain hidden.

The influence of New Communication paradigm. Powered by the web 2.0 technologies, the New Communication paradigm denotes uncontrolled communication patterns between company customers, partners, and employees. In particular, in business to client (B2C) relationship consumers are influencing all aspects of other consumer behavior and finally they are reducing their reliance on advertising as a source of information to guide their purchase decision-making [8]. Moreover, the New Communication paradigm enables faster information circulation. Hence, the environment becomes more dynamic and willing to change.

The discussion in Introduction and above-reflected problems show that it is essential to know when social software that supports mediation of opinions between people is needed in business processes as well as how to introduce it in business processes. To meet these needs at least partly the Iterative and Incremental Social Software integration - IISSI, approach is proposed in this paper. The approach is developed as a complex of diverse methodologies and techniques applied and aimed to organize and guide the integration of social software into business processes. IISSI enables participatory and social enactment improvements of BPM. By applying IISSI approach a business process becomes a socially enhanced business process (see Fig. 1). By social enhancement we mean any improvement provided by social software. Social software can be understood as any internet-based collaboration and communication tool. Socially enhanced business process is a business process enriched by one or many social enhancements. In general, the goal of IISSI approach is to provide simple supporting guidelines for successful social software integration into business processes. IISSI is relatively easy applicable as it incorporates iterative and incremental development (IID), BPM and social software related methods and concepts within a guided development flow.

The approach is based on IID that is a process that builds a system by a gradual increase in features during self-contained cycles of analysis, design, development and testing [12], [13]. IID roots in the work of Walter Shewhart, a quality expert at Bell Labs, in which he proposed to organize quality improvements in series of "plan-do-study-act" (PDSA) cycles [14]. In 1970's the idea was further developed by Harlan Mills, IBM Federal System Division employee, who proposed iterative development model in his work "Top-Down Programming in Large Systems". In the beginning of 1980's Gerald Weinberg published his book "Adaptive Programming: The New Religion" in which he articulates the main idea of IID – "build in small increments with feedback cycles involving the customer reach" [15].

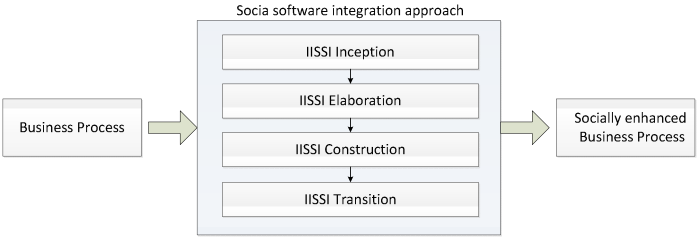


Fig. 1. IISSI approach

The IISSI approach includes four IID Unified Process phases: inception, elaboration, construction, and transition (see Fig. 1). The inception and elaboration phases constitute problem analysis and requirement gathering tasks. The construction and transition phases include iterative development cycles as well as incremental deliveries and deployment tasks. The phases are exposed with 11 IISSI tasks which are supported by several methods and concepts that were found to be helpful for fulfilling these tasks (see Fig. 2). Each IISSI phase and tasks together with the recommended methods and concepts are discussed in the remainder of the paper.

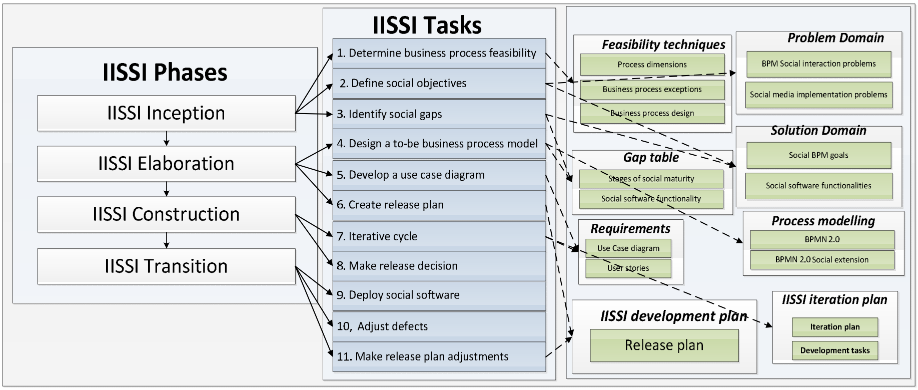


Fig. 2. Overview of IISSI approach (phases, tasks, methods and concepts)

3 Inception

The inception phase of the IISSI approach includes the following tasks (see Fig. 2 and Fig. 3):

- Determine the business process feasibility
- Define social objectives
- Identify social gaps

First two processes help to identify whether there is a need to integrate social software into the business process. There are the following cases when integration of social software might be cancelled:

- The targeted business process is not feasible for social enhancements
- Social objectives do not correspond to the BPM social interaction problems

Thus, inception phase includes two escape mechanisms which cancel the social software integration work (see Fig. 3).

Identify social gaps task is executed in order to determine the IISSI scope. The solution domain serves as a support reference model for the task. The IISSI high level scope is reflected in the gap table. The gap table is used further in the elaboration phase of IISSI approach.

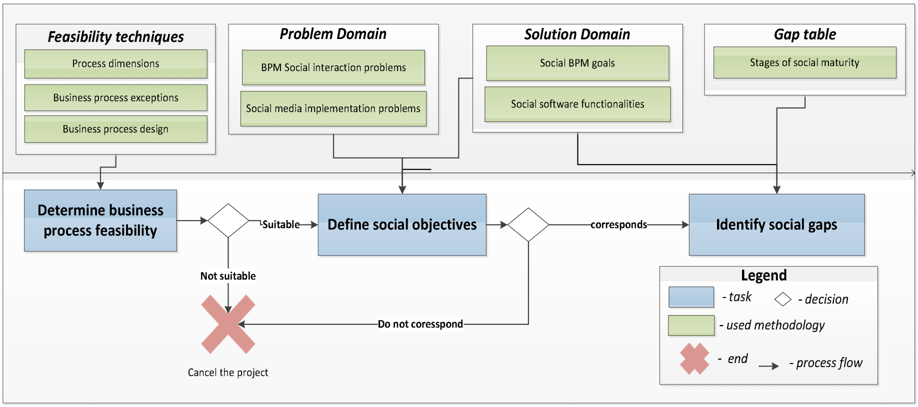


Fig. 3. IISSI approach: Inception phase

In ***Determine business process feasibility***, several methods can be used in order to understand the need of social software integration inside the business process, such as process dimensions, business process exceptions, and business process design. By applying listed methods the decision on proceeding or not-proceeding further with the social software integration should be made. If the feasibility analysis methods discover business process unsuitability to the proposed changes, it is advised to cancel the IISSI work. To apply these methods, usually meetings are organized where business process stakeholders discuss the matter.

When applying *Process dimensions*, the business process execution environment can be analyzed from two different dimensions: the level of role automation and the level of process control automation. The less automated business process (less supported by information technology solutions), the more tasks and flows are guided by actor tacit knowledge [3]. It is believed that less automated business processes are potential candidates for social software integration.

Process implementation environment dimensions model defines the degree for role and process control automation in the business process [3]. The degree metric is exposed as low, moderate or high [3]. The signal for the potential need of social software integration is the low level of role automation and process controls automation.

Business process design method is the way to analyze the business process through observing its design. Two main approaches can be mentioned [6], [16] - *assembly line* and *work station*. Four basic questions can be considered. How much predetermined is the work order? More prescribed – assembly line, less prescribed – work station. Is human communication and collaboration encouraged in the process? Rather not encouraged – assembly line, rather encouraged – work station. What is the level of customization for the end product? Low customization refers to assembly line, high customization – to work station. How high is the process frequency? Rather high frequency – assembly line, rather low frequency– work station. If answers to the most of the questions show business process more similar to work station design, then a potential basis for social software is identified.

Regarding *Business process exceptions*, the aspect to consider is the regularity of process exceptions occurring when performing the business process. Business process exceptions are not often prescribed and lead to the unstructured, spontaneous communications and problem solving [16]. Therefore, potential social software integration need can be identified. The following exception attributes can be considered:

- The occurrence rate
- The level of impact

Define Social Objectives: the social objective is the way how organization can formulate the ends that any potential social software integration enhancement should achieve. It is understood that, in general, the social software solves the motivational business process problems defined in the problem domain. Therefore, the social objectives should correspond to these problem descriptions. If defined social objectives are not suitable to the motivational problems, it is advised to cancel the IISSI work.

The IISSI approach defines *Problem domain* as the environment where the solution will be implemented (see Fig. 4).

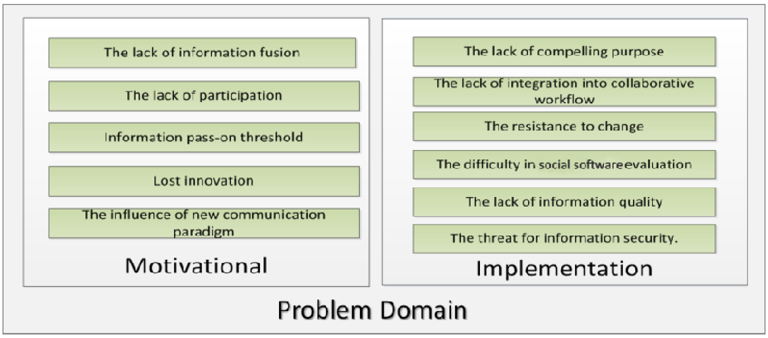


Fig. 4. IISSI Problem Domain

As shown in Fig. 4, the IISSI problem domain includes the following problem groups: (1) motivational – business process social interaction problems and (2) implementation – social software implementation problems.

Solution domain (See Fig. 5) defines the abstract environment where the solution is developed [17]. The IISSI solution domain can be used for defining social objectives and filling the gap table. The IISSI solution domain includes the following concepts:

- *Social business process goals* that concern the question, what potentially can be done in order to improve business process execution. For instance, the goal "to disseminate knowledge" for knowledge sharing can be defined in order to improve business process execution or the goal "establish social feedback" can be introduced to receive feedback from a broader set of actors with the aim to evaluate business process quality.
- *Social software* that concerns the question, how exactly is the social enhancement achieved. For instance, using formal messaging the message exchange from one author to one or several recipients is supported so that the decision distribution environment can be set for making decisions visible to the selected actors or actor groups.

To support the analysis, IISSI approach provides dedicated tables that show when specific goals can be stated, what types of the social software can help to achieve them, and what are known use cases for particular software types.

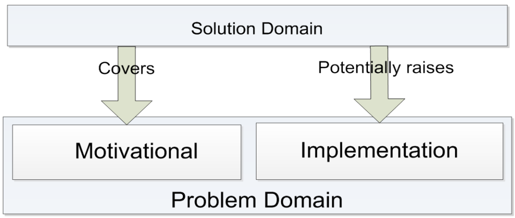


Fig. 5. Relationships between problem and solution domains

The solution domain is aimed to solve the motivational problems of the problem domain. At the same time, it potentially raises the implementation problems.

Gap table is designed to constitute the high level functionality scope of IISSI work. The rows of the gap table are filled until the project stakeholders decide that defined social objectives are fully supported. The gap table concerns the existing social software functionality and to be implemented social software functionality. The gap table is used further in the elaboration phase. The gap table's meta-information is provided in Table 1.

There are mandatory and non-mandatory factors in the gap table. The IISSI approach provides the means for filling the table using different additional tables on issues discussed in this section. One of the examples is the table about stages of social maturity [16] (see Table 2). This table helps to fill the gap table concerning two factors, namely, Social maturity stage (as-is) and Social maturity stage (to-be).

4 Elaboration

In the elaboration phase the social software integration project starts to take the form. The project plan and the architectural foundation are established. Functional requirements are described and transformed into the use-case model and user stories [18], [19]. At the end of the elaboration phase there is achieved a clear understanding of the functionality scope and technological foundation that will support the most essential use cases.

The elaboration phase of the IISSI approach includes the following tasks (see Fig. 6):

- Design a to-be business process model
- Develop a use case diagram
- Create a release plan

Table 1. Meta-information for Gap table

Factor name	*Mandatory?*	*Options*	*Description*
Unique Identifier	Yes	Number	A unique identifier of the gap.
Social objective	Yes	Textual information	Social objectives derived from the task "Define social objectives".
Social software	Yes	Textual information	Social software from the solution domain is depicted.
New	Yes	Yes / No	Identifies whether the social software functionality is new within the business process.
Business process task	No	Textual information	Identifies business process task(s) where social software is used.
Gap type	No	Quantitative, Qualitative, Redundant.	Quantitative – the social software was not used before in the business process. Qualitative – notates the needed improvement for the existing social software. Redundant – the social software is already present in the business process task (e.g., social software is picked twice for the same business process task).
Social maturity stage (as-is)	No	Stage 1, Stage 2, Stage 3.	Stage 1 - the social software is outside business process workflow and does not correspond directly to any of the business process tasks. Stage 2 - the social software is used as a process exception handler for a business process task or it is needed to change working environment in order to access the tool when executing the business process task. Stage 3 - social software is used in the business process task execution and it is accessible within one technical solution.
Social maturity stage (to-be)	Yes	Stage 1, Stage 2, Stage 3.	The same as for "Social maturity stage (as-is)".
Description	Yes	Textual information	Informal description of the intended functionality.

In the elaboration phase the defined gaps in the gap table are considered and BPMN 2.0 (recommended) notation is used for designing to-be business process model. Slightly modified use case diagram and user stories are used to define social enhancement functionalities. At the end, the release plan is created. The decision is made whether it is feasible to proceed further to the construction phase.

Design a to-be Business Process Model. The architectural foundation for the IISSI approach is the business process model. The non-executable business process model is suggested to be developed in BPMN 2.0 notation. The business process model exposes social software functionalities taken from the gap table. The social software functionality placement is the following:

- As a support functionality outside the process model – social maturity stage 1
- As a process model exception handler - social maturity stage 2
- As a process model task - social maturity stage 3

Table 2. Stages of social maturity derived from [16]

Stage	*Title*	*Description*
Stage 1	Isolated communities, outside the workflow	Bottom-up contributions of stakeholders (employees, clients, other actors) within communities free from existing organizational workflows. Low cost solution that does not require additional integrations.
Stage 2	Communities outside the flow, supporting a traditional process	Bottom-up contributions of stakeholders support a particular workflow that remains unchanged. These contributions compliment the workflow exceptions or change requests with necessary tacit knowledge. The users need to change between unstructured and structured environments which are not integrated into the common solution.
Stage 3	Socialized process	Unstructured and structured work is placed within one technical solution. Bottom-up contributions such as comments, ranks, shares, and status are integrated within the workflow. However, these collaborative workflows remain as isolated workflows. Supported technologies do not communicate with each other.
Stage 4	Integration of socialized processes	Collaboration services are unified together serving the multitude of social interactions. These processes are producing updates which are visible for every individual through a status update flow. Users are benefiting as they stay up-to date to the latest changes of processes. Processes are benefiting as users can instantly contribute (provide feedback, comment, or add information) to the process value chain.

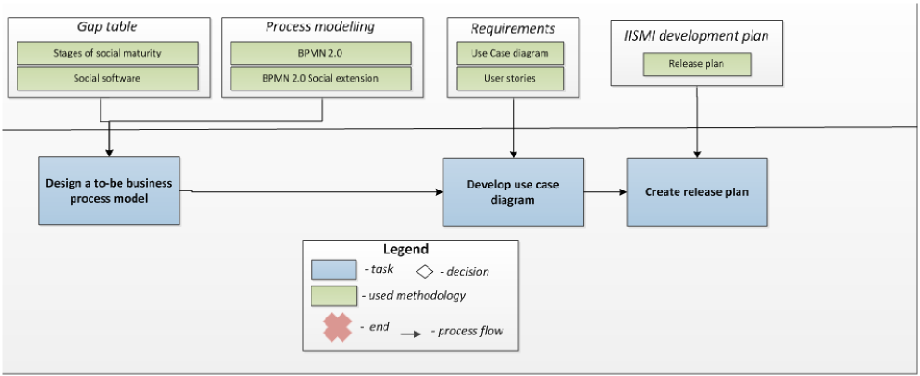


Fig. 6. IISSI approach: Elaboration phase

Social software functionality of social maturity stage 1 does not change the flow of the business process model. However, there is a need to incorporate social software functionalities within business process model in social maturity stages 2 and 3 (see Fig. 7). This is done by using BPMN 2.0 social extension notation [20].

Develop Use Case Diagram. Use case diagram is a simplistic representation of user's interaction with the system. In the IISSI approach the use case diagram is used with bit different user and system notations. The users of the system are business process model actors. The system is the business process. The actor and the business process interactions are supported by social software functionalities. The use case diagram should be enriched with the user stories that describe interactions in the more elaborated way.

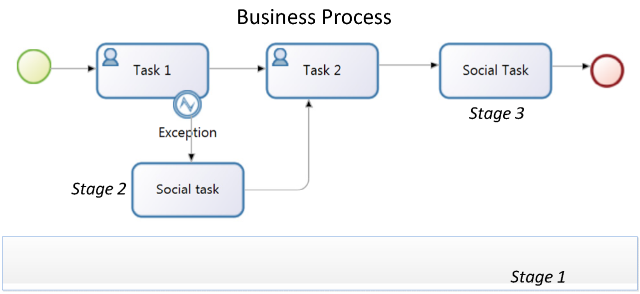


Fig. 7. Social software functionalities in the business process according to social maturity stages (Stage 1, Stage 2, Stage 3)

Create Release Plan. In IISSI approach the release plan consists of several release and iteration estimations. The primary goal of the release plan is to estimate realistically what user stories of the social software functionalities will be implemented in the first release, what will be implemented in the second release, and so forth [21]. The client decides which are the most important user stories that should be the first ones to be developed [22]. The project plan includes the following information:

- Number of releases with corresponding release dates
- Available resources for each release
- Number of iterations for each release – preferably iteration length should not exceed 3 week time period
- User stories with estimations included– the most important user stories come to the first releases

The release plan can be changed if the development speed in the iterations changes drastically. Then a release meeting is scheduled and the new release plan is created [23].

5 Construction

In the construction phase the high-level requirements are further analyzed. As the result, the detailed development plan (*the iteration plan*) is derived and the actual implementation can begin. The development is organized by executing several iterations that are part of a particular release plan (see Fig. 8). The release here is a stable executable version of the social software support which can be deployed into the operational environment [24]. Each release constitutes an increment – an increase of social enhancements in the targeted business process.

Develop Iteration Plan. The iteration can be understood as “a self-contained mini-projects with a well-defined outcome: a stable integrated and tested outcome” [24]. The notion of the "mini-project" helps development team and other stakeholders to focus on the most essential functionality first. Usually the iteration length is in the space of 1 to 3 weeks [25]. The iteration is a unique activity. Therefore, the iteration requires the iteration plan [24]. The iteration plan contains the list of user stories and corresponding development tasks [26].

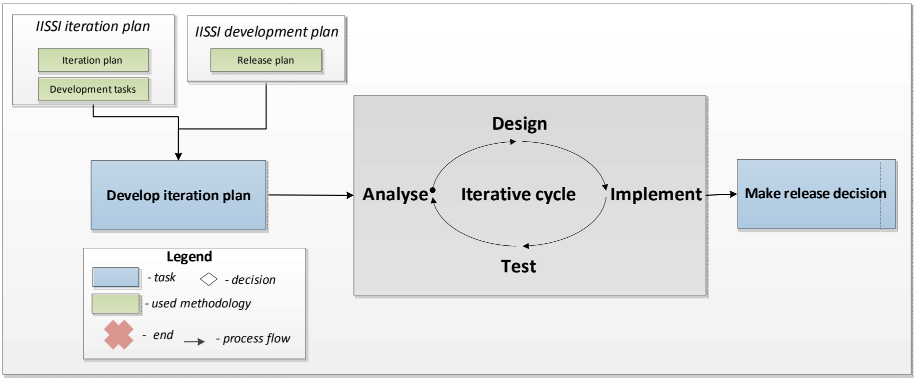


Fig. 8. IISSI approach: Construction phase

The iterative cycle of the IISSI consists of the analysis, design, implementation and test stages. Each iteration should result in better understanding of requirements (user stories), technical architecture, increased development knowledge and more complete implementation.

Make Release Decision. At the end of the last iteration in the release the decision over closing the release is made. If it is positive, the transition phase is launched. The next iteration cycle of the next release starts.

6 Transition

Transition is the final phase of the social software integration project done by IISSI approach (See Fig. 9).

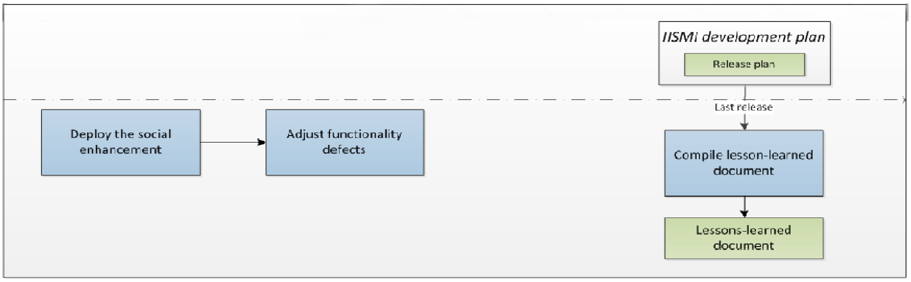


Fig. 9. ISSI approach: Transition phase

The transition phase includes the functionality deployment to the production operating environment. Once the functionality is delivered, the additional defects are corrected; functionalities are adjusted to the specific user needs [19].

The transition phase of the IISSI includes the following tasks:

- Deploy the social software functionality to the production operating environment

- Adjust functionality defects
- Compile lesson-learned list

Deploy the social software functionality task incorporates all the activities needed for making social enhancement available for use.

Adjust social software functionality defects refers to making needed changes when some problems with implemented social extensions are found.

Make Release Plan Adjustments. The release plan adjustments are made in order to accommodate the user stories initialized from gathered feedback. Additional iteration can be planned or user stories can be incorporated within the scope of current release plan.

7 IISSI Application Example – Foreign Studies Planning

IISSI approach was applied for the foreign studies planning (FSP) process at Riga Technical University. FSP is an annual process in which the list of courses to be delivered in English is compiled for the next study year. IISSI approach was applied in order to decrease overall time spent in doing FSP work.

Main methods and concepts suggested by IISSI approach and discussed in previous sections are amalgamated in Table 3. The methods used in the application example are depicted in Fig. 10.

Table 3. Methods and concepts suggested by IISSI approach

Nr.	*Title*	*IISMI Phase*	*Usage*
1.	Business Process Model and Notation (BPMN)	Elaboration	Used to design the to-be business process model
2.	BPMN 2.0 social extension	Elaboration	Used to annotate the social media integration in the to-be business process
3.	Social BPM goals	Inception	Along with social software constitutes the solution domain
4.	Social software	Inception	Along with Social BPM goals constitutes the solution domain. Used in defining the gap table
5.	BPM Social interaction problems	Inception	Along with social media implementation problems constitute the problem domain
6.	Social media implementation problems	Inception	Along with BPM Social interaction problems constitute the problem domain
7.	Stages of social maturity	Inception, Elaboration	Used in defining the gap table and designing to-be business process model
8.	Release plan	Elaboration	Used to plan overall IISMI development
9.	Iteration plan	Construction	Used to plan IISMI iteration development
10.	Process dimensions	Inception	Used to plan IISMI iteration development
11.	Business process exceptions	Inception	Used to determine business process for social media integration
12.	Business process design	Inception	Used to determine business process feasibility for social media integration
13.	Development task	Construction	Used for transforming user stories into specific requirement text

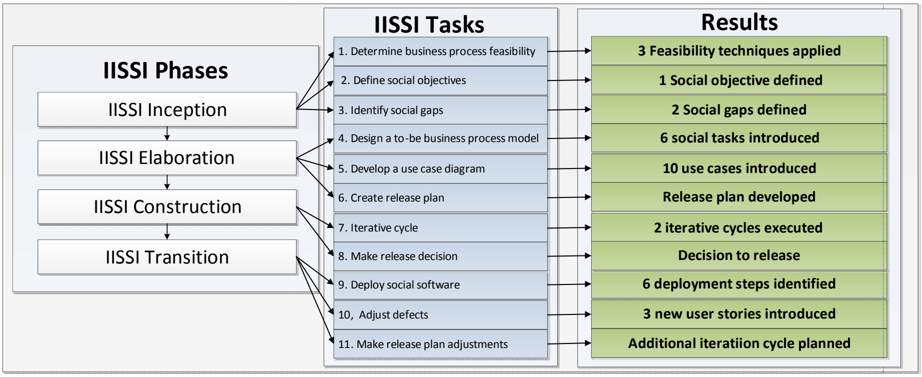


Fig. 10. Overview of application of IISSI approach

The information pass-on threshold and the lack of the information fusion as BPM social interaction problems were identified within the as-is FSP business process (see Section 2). The information pass-on threshold is denoted by "1" and the lack of the information fusion is denoted by "2" in Fig. 11.

As the result of IISSI inception and elaboration phases, the Wiki social software was proposed in order to handle the updates of the list of English courses within the centralized repository. Five process tasks were removed from as-is FSP business process model and 6 new social tasks were introduced into to-be FSP business process model (see Fig. 12).

As a result of IISSI construction and transition phases, the Wiki "MediaWiki" social software was configured. The overall time spent in preparing course list in the to-be business process was decreased as the course lists were not duplicated and aggregated (the lack of information fusion) and were not sent back and forth between process actors (information pass-on threshold).

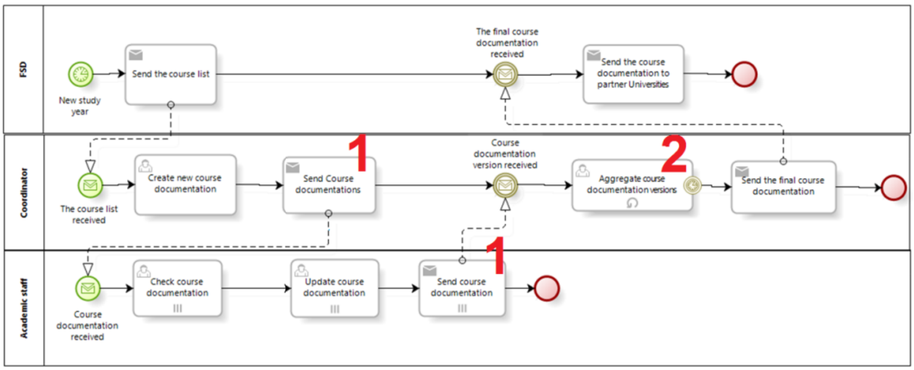


Fig. 11. As-is FSP business process

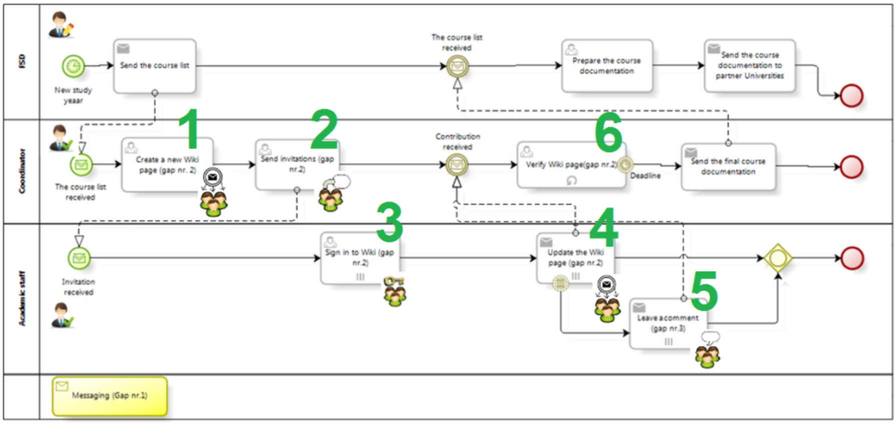


Fig. 12. To-be FSP business process

Thus the application of IISSI approach gradually successfully resulted in 6 socialized tasks in the FSP business process and the Wiki software was implemented at social maturity stages 2 and 3.

8 Conclusions

The IISSI approach is aimed at overcoming social integration problems in BPM and was tested on foreign studies planning process at the university. IISSI approach organizes different methods and concepts around the well known iterative and incremental development approach. It provides the following contributions to social software integration in BPM:

- Proposes logical structure for including BPM and social software related methods and concepts in a guided development flow. It is helpful as at the moment there is a lack of supporting guidelines for incorporating social software into the business process development.
- Encourages the domain learning and contributes to continuous quality improvements by using iterative approach. It is helpful as detailed social software integration requirements can hardly be established at the beginning of the social software integration effort.
- Suggests concrete escape mechanisms that stop the process of unfeasible social software integration at early development stages. Social extension is not always beneficial and can be excessive in cases where social interactions do not require additional management. Such cases are common within highly automated and repetitive processes.
- Shows how business process modeling can be used for annotating as-is and to-be processes and identifying activities where social software functionality can be incorporated.

Future work concerns further testing of the approach with respect to more social interaction problems and development of metrics, that can help to assess social software integration benefits in complex business processes.

References

1. Hammer, M.: Reengineering Work: Don't Automate. Obliterate. Harvard Bus. Rev., 104–112 (1990)
2. Brocke, J.V., Rosemann, M.: Handbook on Business Process Management 1. Springer, Heidelberg (2010)
3. Ward-Dutton, N., Macehiter, N.: Business process management: a holistic view (2005)
4. Cleveland, S.: 6 Benefits of BPM., http://www.ebizq.net/blogs/bpm_business/2011/10/6_benefits_of_bpm.php
5. Harmon, P., Wolf, C.: The State of Business Process Management (2008)
6. Erol, S., Granitzer, M., Happ, S., Jantunen, S., Jennings, B., Johannesson, P., Koschmider, A., Nurcan, S., Rossi, D., Schmidt, R.: Combining BPM and social software: contradiction or chance? J. Softw. Maint. Evol. Res. Pract. 22, 449–476 (2010)
7. Schmidt, R., Nurcan, S.: BPM and Social Software. In: Ardagna, D., Mecella, M., Yang, J. (eds.) BPM 2008 Workshops. LNBIP, vol. 17, pp. 649–658. Springer, Heidelberg (2009)
8. Mangold, W.G., Faulds, D.J.: Social Media: The new hybrid element of the promotion mix. Bus. Horiz. 52, 357–365 (2009)
9. Surowiecki, J.: The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations (2004)
10. Granovetter, M.: The Strength of Weak Ties: A Network Theory Revisited. Sociol. Theory 1, 201 (1983)
11. Merz, M.A., He, Y., Vargo, S.L.: The evolving brand logic: a service-dominant logic perspective. J. Acad. Mark. Sci. 37, 328–344 (2009)
12. MSDN: Definition: Iterative and Incremental Development (2007)
13. Techopedia.: Iterative and Incremental Development (2010)
14. Larman, C., Basili, V.: Iterative and Incremental Development: A Brief History. Computer 36, 47–56 (2003)
15. Weinberg, G.: Adaptive Programming: The New Religion (1980)
16. Social Business: Cooperation vs Competition. Harvard Bus. Rev. 6 (2012)
17. Cunningham, W.: Solution Domain (2010)
18. ZuehlkeEngineering: 3C05: Unified Software Development Process, http://www.zuehlke.com
19. Bigelow, D.: Phases and Iterations, http://www.bigelow.ch/Process/Pr_Phases.aspx
20. Brambilla, M., Fraternali, P., Vaca Ruiz, C.K.: Combining social web and BPM for improving enterprise performances. In: Proceedings of the 21st International Conference Companion on World Wide Web, WWW 2012 Companion, p. 223. ACM Press, New York (2012)
21. Waters, K.: Agile Release Planning (2008)
22. Wells, D.: Release Planning, http://www.extremeprogramming.org/rules/planninggame.html
23. Wells, D.: Release Plan, http://www.extremeprogramming.org/rules/commit.html
24. Spence, I., Bittner, K.: Managing Iterative Software Development Projects (2006)
25. Wells, D.: Iterative Development, http://www.extremeprogramming.org/rules/iterative.html
26. Wells, D.: Iteration Planning, http://www.extremeprogramming.org/rules/

Generic Components for BPMN Specifications

Peter Forbrig

University of Rostock, Department of Computer Science, Albert-Einstein-Str. 22
18059 Rostock, Germany
Peter.Forbrig@uni-rostock.de

Abstract. The paper introduces generic components to BPMN that reduce modelling efforts and ensure higher quality of specifications. The idea is transferred from task models to business process models. Some of the introduced components can be considered as patterns. Therefore, the paper discusses different views on patterns and their relations to reusable generic component. Additionally, different strategies for adaptation like design time adaptation and runtime adaptation of components are discussed. Parameter substitution is supported best by tools. Requirements for such support are finally collected and discussed.

Keywords: Business process modelling, BPMN, generic components, patterns.

1 Introduction

There is no doubt anymore that modelling of domains provides a lot of advantages. It is an essential approach for understanding the interactions between different components and stakeholders. Business process modeling focuses on understanding, documenting and improving of workflows. It is becoming a high priority for analysts. “BPM has enabled many organisations to eradicate the chaos of historically evolved business procedures and made them more structured, transparent and standardized.” [19]

There are different notations for business processes like UML activity diagrams [24], event process chains [17] or BPMN [1]. Meanwhile a lot of tools exist that allow the specification of these models with their editors. However, it is still challenging to create models for enterprises. This relies on the fact that such models are large. Starting from scratch is very time consuming. Starting with specified models hast the problem that sometimes remodeling is not performed completely and the result becomes wrong.

Reuse of existing parts of models would be very helpful. This is can in some way be done by using workflow patterns (see e.g. [15], [26]). These patterns provide solutions for reoccurring problems. However, the patterns are on a very low level of abstraction. They are very helpful in specifying details like sequence, parallel split or cancel case. However, they do not help reusing business processes like buying or selling something. Reference models like presented in [18] are more helpful for this purpose. Such models provide a standard solution for standard tasks for enterprises. However, the reuse of such model is not supported by tools. Humans have to adapt

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 202–216, 2014.

the provided solutions. This is still time consuming and costs a lot of human capital. The main problem is the missing support by refactoring the existing models like renaming of processes and artifacts. Adaptation of workflows lacks support as well. To provide this support a concept of generic components is discussed that will allow the reuse and adaptation of models.

The paper provides the concept of generic components for BPMN. It is structured as follows. First a short example will motivate and discuss the principles of the suggested approach. Afterwards, some more complex examples will be presented and requirements for tool support will be provided. In paragraph three there will be the discussion of related work and finally there will be a summary and an outlook.

2 Reuse of Models in BPMN

During the last twenty years it became obvious that modeling is an important aspect for software development and organizing the work people have to perform. A lot of specification languages appeared. We will focus on BPMN in this paper. For requirements analysis and brain storming writing on paper is a good way of modeling. However, maintaining the specifications for a longer period asks for tool support. There are already a lot of tools available that allow combining basic elements of the corresponding language. Nevertheless, the specifying with such tools is still challenging, time consuming and error prone.

Therefore, reuse of already developed specifications is very helpful. However, at the moment the reuse of workflow specification is often restricted to a specific context. We will discuss some ideas how this can be changed. This paper will suggest an extended notation for BPMN that allows the reuse of generic components.

2.1 Introducing Example

Let us first have a look at an example workflow that has the potential to be used in different contexts. It was used on a website to motivate the notation of BPMN and the usage of the corresponding tools.

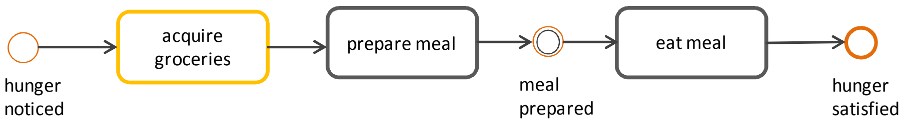


Fig. 1. BPMN specification for hungry people (taken from [1])

A person recognizes that he/she is hungry. What can be done? According to the provided workflow one has to acquire groceries and prepare a meal afterwards. When the preparation is finished there is a prepared meal and on can start to eat. After eating is finished the person is not hungry anymore.

The workflow specification has really the potential to be reused in different contexts. For this purpose it makes sense to provide an abstraction and introduce a parent process called “solve problem hungry”. This process node can be used within different workflow specifications.

There is only one problem. What happens when the specification of the sub-process changes? One could want to have always a reference to the newest version or one would like to want the state of the specification from that moment the parent node was introduced into a workflow specification.

There are good reasons for both strategies. It depends on the context which one is preferable. Therefore, in editors there has to be an option for developers to select the strategy they want to apply. In case there has always to be used the most updated version of the sub-workflow specification the reference to this specification has to be stored. Otherwise, a copy of the sub-workflow specification has to be stored. Fig. 2 provides the notation for both cases in BPMN.

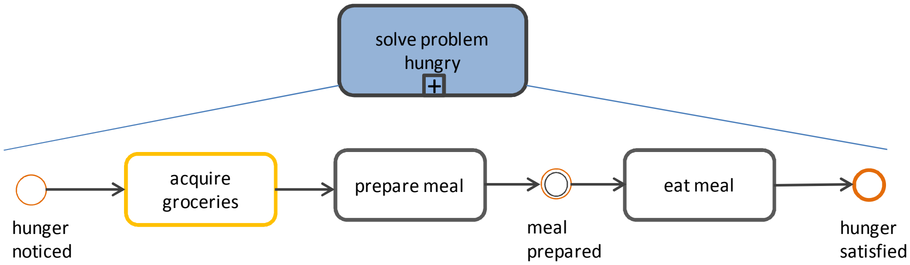


Fig. 2. BPMN specification for hungry people with parent node

With the introduced parent process reuse of the workflow specification is easy possible. However, applications are very much restricted to a specific context. In case one wants to specify a concrete meal like a pizza this would be possible only by copying the specification and editing it. However, this is time consuming, error prone and makes it difficult to realize improved updates from the original specification.

Additionally, it might make sense to provide a workflow specification that is applicable for thirsty people as well. There really seems to be a similar solution for this situation. Thirsty people also have to acquire something. In their case it can be called ingredients. However, the verb in sub-process three has to be generalized to consume because it has to fit to meals and drinks.

A generic specification would help to solve the mentioned problems. For the presented workflow specification from Fig. 2 it might make sense to provide three parameters. The first one (P1) for different problems like hunger or thirst that are recognized. The second one (P2) tor things that have to be acquired and finally (P3) for the object that is consumed. In this way it will be possible to have the workflow specified for meal, pizza or sushi.

According to these findings the workflow specification has to be rewritten. The following figure provides the resulting specification.

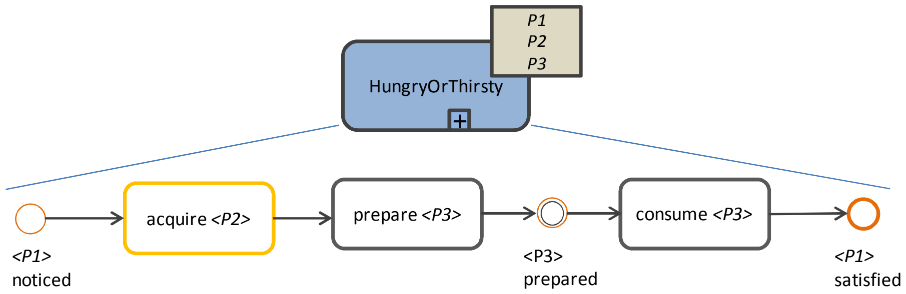


Fig. 3. Abstracted generic component for hungry or thirsty people

Parameter P1 can have the value hunger or thirst. For P2 groceries or ingredients are possible values. P3 can have the values meal, drink, pizza, Manhattan or something else. The names of the parameters can be selected more domain specific. Because of brevity short names were used in this first example. (Domain specific names help to understand the instantiation process and make instances more readable. They will be used in the forthcoming examples.)

To instantiate a generic component, values have to be assigned to the formal generic parameters. Such instances can be written in the following way.

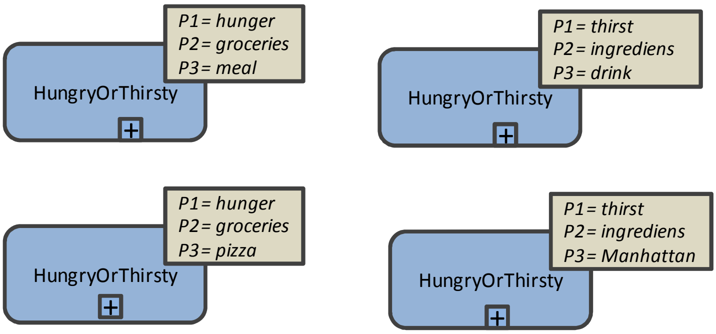


Fig. 4. Instances of generic components

Within the workflow specification all appearances of formal parameters are substituted by the corresponding assigned values. The upper left instance of Fig. 4 results in the specification already presented in Fig. 1. The other instances deliver corresponding specifications.

Hopefully, the provided example was already convincing that the concept of generic components is useful for workflow specifications. However, a more complex example might be helpful as well. For this reason an example was selected that is provided together with the definition of BPMN by the OMG [1].

2.2 More Complex Example

The example provided by OMG [3] is related to our introducing example. It is again about hunger and eating. It provides a workflow specification for ordering and delivering a pizza. The specification of Fig. 5 starts with the situation that a pizza customer is hungry for a pizza. After selecting a pizza a corresponding order is placed. A pizza vendor takes the order, bakes the pizza and delivers it. Additionally, some details are specified if the pizza is not delivered in time and how the payment will be performed.

One can consider this specification as reference model for all food delivering services. However, the workflow specification has to be refactured to a certain extend to support the adaptation to different contexts.

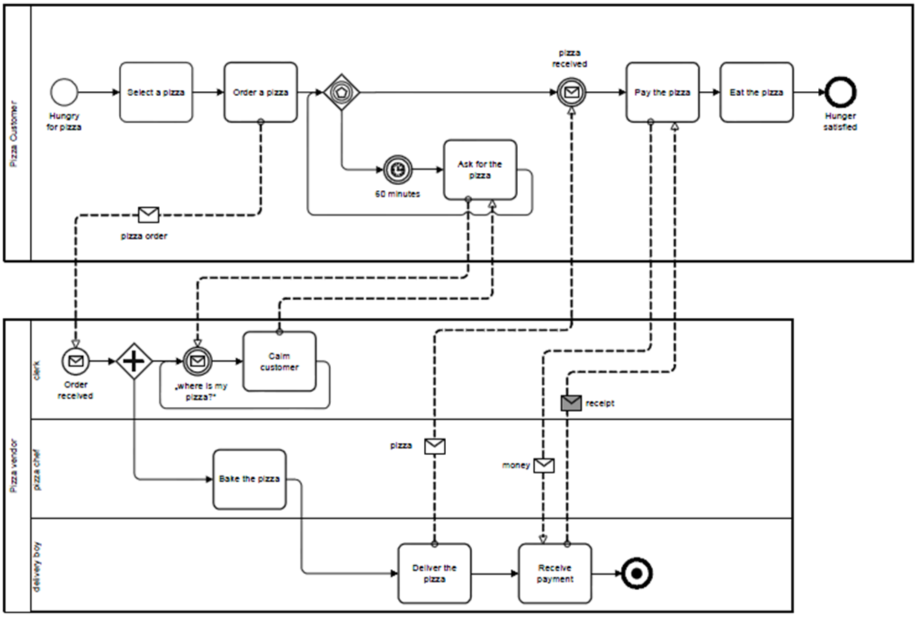


Fig. 5. BPMN specification for ordering and delivering pizza (Figure 5.2. from [3])

Again this process can be generalized by renaming some elements and by introducing parameters. Following the already discussed approach the workflow specification can be adapted for ordering and delivering sandwiches.

Generalizing it even more it would be possible to use this process specification for ordering calendars as well.

In the following specification the item that is ordered is represented by parameter <sthg>, which stays for something. The term “Hungry for Pizza” is generalized to “Keen on <sthg>”. Correspondingly “Hunger satisfied” is replaced by “Satisfied with <sthg>.

Additionally, “pizza customer”, “pizza vendor” and “pizza chief” are replaced by “<sthg> customer”, “<sthg> vendor” and “<sthg> creator”. Together with some further parameterizations these changes result in the following specification of a generic component.

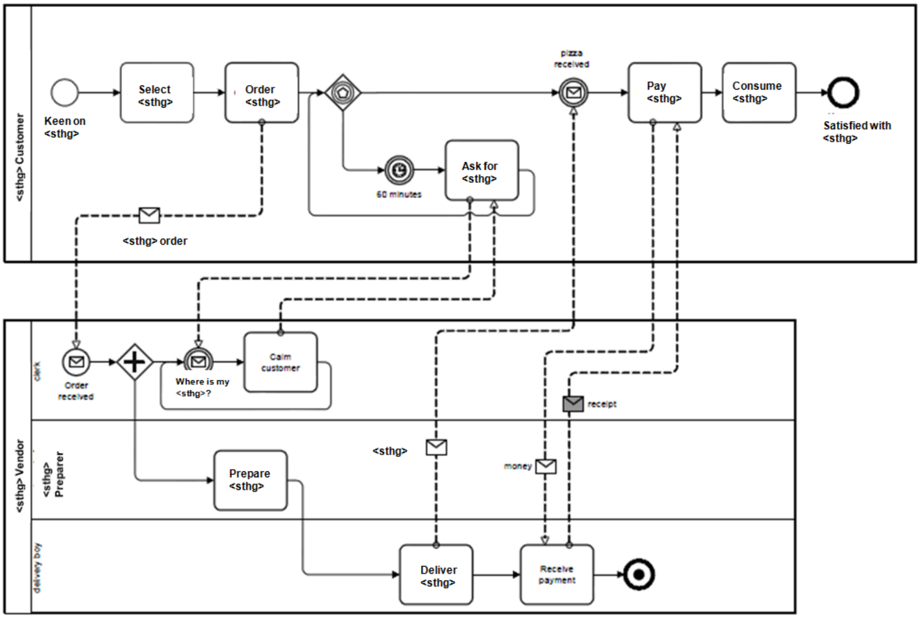


Fig. 6. Modified and parameterized BPMN specification of Fig 5

The presented specification can be represented by a generic component and by its instances. The following figure provides the graphical representation of these components.

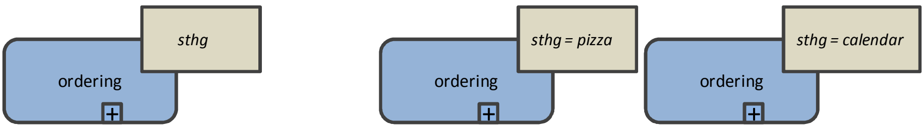


Fig. 7. Generic component and two of its instances

Looking at the workflow specification of the instance with value calendar for the generic parameter it becomes obvious that 60 minutes might not be appropriate for asking for the result. Indeed, it makes sense to introduce a further generic parameter for the time that is appropriate for complains.

Additionally, one can recognize that there is a sub-process “Consume <sthg>” that might not be appropriate in all cases. It might not be necessary for a calendar. There is a need for a third parameter. Its value can be used to execute a design time decision.

The following figure presents the corresponding specification. If the parameter consume has the value “no” there is no sub-process for consumption. However, if the value is “yes” the corresponding sub-process is still part of the specification. In this way specifications can be adapted due to the provided values of parameters. Fig. 8 highlights this fact by focusing on the essential part of the specification.

Design time decision means in this context that a tool supporting the idea of generic components of BPMN should be able to perform the described transformation of the workflow specification according to the parameters of a component during instantiation.

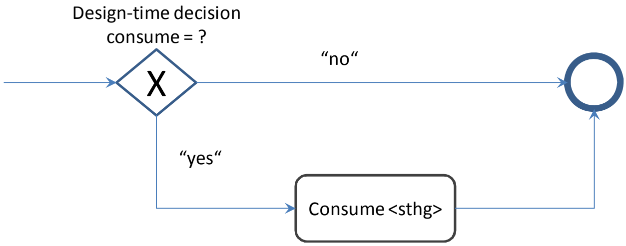


Fig. 8. Specification of a design-time decision

BPMN normally asks for a task in conjunction with a XOR-gateway. The task is executed during run-time and its result is used to activate the forthcoming task or tasks. For our purpose such task is not necessary because the decision is taken based on the value of parameters during. In this way the workflow specification is adapted at design-time.

In case that the component specification is extended by two further parameters the component and its instances would look like presented in the following figure.

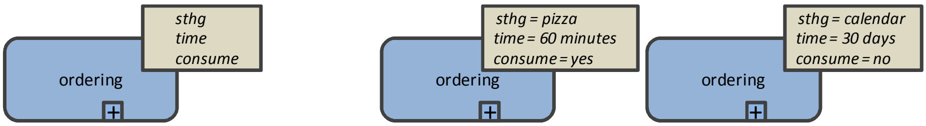


Fig. 9. Generic component with three parameters and two of its instances

The business process is adapted according to the value of the parameters. In case of a pizza there is a process step “Consume pizza”, whereas for calendars this process step does not exist. This is at least true according to the in Fig. 8 provided business process specification.

2.3 Run-Time versus Design-Time Instantiation and Adaptation

The component of Fig. 9 has three parameters that are pretended to be substituted during design time. However, for parameters like consume it might be possible that their value is still not available at design time.

Instead of “yes” or “no” there might still be a question mark “?” during instantiation. In this case the needed value does not yet exist. The decision for selecting the appropriate path has to be postponed to runtime.

For this postponed decision there can still be different strategies. The first one would postpone the decision only to that moment where the execution of the component starts. In this way the specification is adapted at start time.

Another strategy would be to postpone the decision until the decision at least has to be made. This kind of strategy can be characterized by the attribute “lazy” or as execution time fixed.

The selection of the appropriate strategy during runtime can be made by the designer of the workflow specification or it can be delegated to the animation or runtime system.

If the parameter substitution is dependent on runtime information this can be performed at start time or execution time. This depends on the current context and can be different in similar situations.

The following strategies for parameter value substitution were identified:

1) At design time
2) At runtime
    a) At start time fixed by the designer
    b) At execution time fixed by the designer
    c) At that time where parameters are available (start or execution)

The identified strategies ask for language and tool support. Regarding BPMN there should be the opportunity for related annotations to decisions like presented in Fig. 8. The annotations could be “Design-time decision”, “Start-time decision”, “Execution-time decision” and “Available-time decision”.

There was a suggestion by an anonymized reviewer to characterize the suggested approach by model layers in a graphical way. Figure 10 is an attempt to follow this advice.

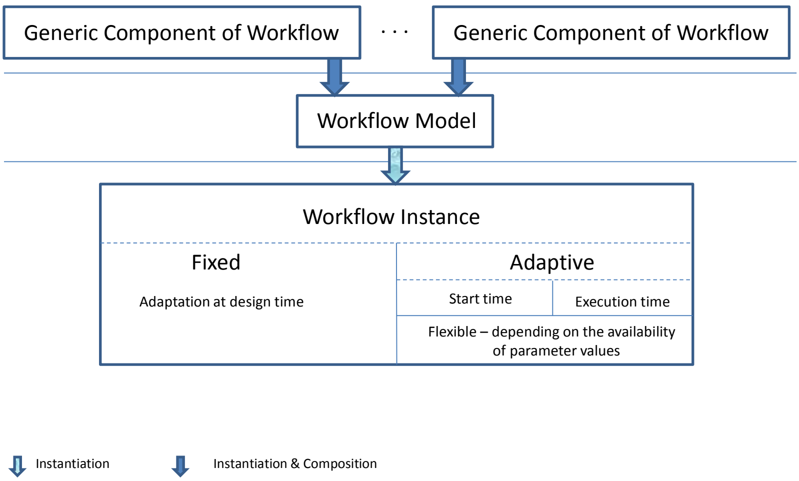


Fig. 10. Specification of a design-time decision

Workflow models can be composed of instances of generic components. Sometimes instantiations are performed beforehand. However, these instantiations can also be postponed.

The model of a workflow is able to instantiate workflows that are fixed in case all parameter of generic components were already instantiated during design time. For still adaptive workflows there are three options. They are instantiated at start time, at that time the instantiation is specified or the specification was flexible depending on the availability of parameters from the context. This can be in one case at start time of the workflow and in the next case during runtime.

Further requirements for tool support will be discussed within the next paragraph.

2.4 Requirements for Tool Support

We already mentioned the necessity of appropriate tool support for using generic components in BPM. This paragraph is intended to summarize the requirements for such tools.

1 Specifications of components must be supported by two options. a) A copy of the instance is computed and stored. b) The instance is computed always while browsing the instance specification.
2 While browsing instances of components the corresponding values of parameters have to be substituted.
3 Design time decisions within specifications have to be executed while instantiating generic components.
4 An optional requirement would be the ability of a tool to animate specifications.

The last requirement is mentioned the first time. It is not related to the provided idea of generic components but comes from the experience with task models. Tools like CTTE (Concurrent-Task Tree Environment) [6] or Hamsters provide such kind of support. The corresponding models can be animated in such a way that the models are interpreted and possible next tasks are computed. One of these tasks can be interactively selected. As a result further interpretations are performed and a new executable task set is computed. Such kind of animation would be possible for BPMN-specifications as well.

In [5] this idea was already used for UML activity diagrams. Animation can proceed if there is only one action that can be performed. If there are more than one the selection has to be done interactively.

3 Related Work

3.1 Generic Components in UML Class Diagrams

Generic Components are well known from programming. Nevertheless, they are rarely used in modeling. An exception is the concept of generic classes in UML.

However, the generic concept is used for single classes only. Nevertheless, it was an inspiration for the suggested extension of BPMN. The following figure presents an example of a class diagram.

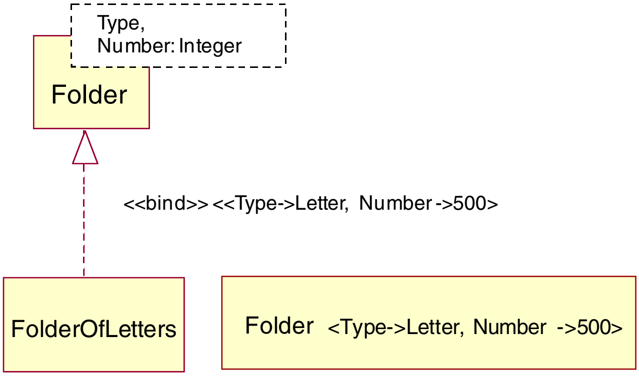


Fig. 11. Generic class folder and two ways of instantiation

Figure 11 provides a class folder with two generic parameters. The first one represents the type of elements the folder manages and the second one the number of elements. Additionally, two ways if instantiation are provide to get a folder of 500 letters. Because of better readability another notation was suggested for BPMN. However, it would be possible to use the UML notation as well if this is considered to be better readable.

It was already mentioned, that UML allows only single classes to be generic. There is no concept for updating whole diagrams according to provided parameters. To the best of our knowledge there still exists no tool that allows the instantiations for class diagrams.

3.2 Design Patterns

However, there is a notation for applying GoF design patterns [8]. Such a transformation of class diagrams by GoF design patterns is somewhat related to the suggested concept of this paper.

There are tools like Together [22] or Rational Rose [13] that provide support in mapping elements of the pattern to elements of a class diagram and the corresponding transformation of this diagram. The mapping function is not restricted to specific elements of the pattern. Therefore every element can be considered as generic parameter. UML provides even a notation for that, which is exemplarily represented by Fig. 12. Unfortunately, most of the existing tools neither support this kind of diagram nor the corresponding transformation. To the best of our knowledge Together seems to provide the best support, even that this is still not optimal.

Nevertheless, there might be the idea of using the same ideas in conjunction with workflow patterns.

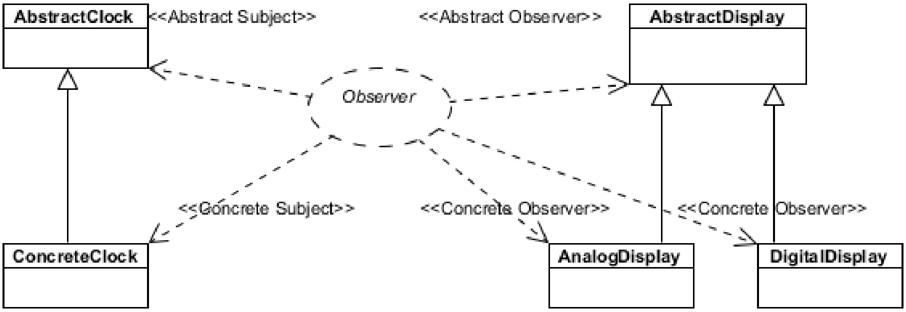


Fig. 12. Pattern Observer and its mapping to a specific class diagram

3.3 Workflow Patterns

Unfortunately, workflow patterns are up to now discussed on a lower level of abstraction only (see e.g. [26] or [29]). They describe solutions for workflows on the level of language features like sequences, alternatives, procedure, metaphors etc. Such abstract solutions and naming conventions like for GoF design patterns do not exist for workflow patterns.

For more complex reuse of knowledge there exist reference specifications of workflows (see e.g. [17]). They are widely used and allow companies to models their models according to the reference specification. However, for such models the generic parameters are missing. In this way formal rules for adaptation are missing. With the suggested approach reference models can be refactored in such a way that opportunities for more precise reuse of knowledge are reached.

3.4 Task Patterns and Generic Task Components

Like workflow specifications task models describe behavioral aspects of humans and systems. It was shown in [5] that task models can be represented as structured UML activity diagrams. Corresponding tool support was demonstrated.

First task patterns were described in [4]. The idea of generic task patterns was the first time suggested in [22]. Corresponding tool support was demonstrated in [12]. However, in both cases the substitution of parameter values was considered to be performed during design time only. The implementation of runtime substitution of parameter values was discussed in [7] the first time for such models. For those patterns the notation of HAMSTERS [9] was used that allows sub-models, procedures and conditional sub-trees.

A case study was performed and the result was that generic task patterns eliminated modeling bugs. It was recognized that the repeated specification of similar workflows did not always result in similar models. Certain details were sometimes forgotten to be modeled. These mistakes were eliminated by using instances of generic components. Additionally, it was possible by using generic task components to reduce the size of models.

The case study was about modeling tasks for a ground segment application that is currently used to monitor and to control the Picard satellite that was launched in 2010 for solar observation. The flat model without using composition mechanisms consisted of 59 tasks.

Sub-models and sub-routines reduced the number of tasks by 24 %. The usage of generic components resulted in a further reduction of 17%. These figures are very much dependent on the kind of models. However, they demonstrate that model specifications can be reduced considerably.

This result makes us hope that the for business process specifications the amount of reduced model specification can even be higher. Because of a missing concept of sub-routines in BPMN the expected reduction of modelling efforts by generic components can be estimated even higher. The reuse of such models saves a lot of efforts. This can already be seen from example presented in Fig. 6.

3.5 Reference Models

Additionally, this is especially based on the fact that in business informatics the idea of so called reference models is wide spread. W. A. Scheer and other authors provide a lot of reference models for companies. We will only refer to those presented by Scheer in his extended Event Process Chain (eEPC) notation. The interested reader is referred to [17] or [18]. On the web page of his business-process tailoring tool [20] it is mentioned:

“One of the greatest challenges for businesses is the dynamic adaptation of business processes to available resources, changing market and business situations.”

This is very important for companies for their business strategies. However, it is also important for workflows, their specifications and the supporting tools. We strongly believe that generic components can support the ongoing dynamic adaptation of business processes in an excellent way. The parameters of the generic components provide an additional control mechanism for adaptation.

3.6 Feature Models

Another approach for the configuration of workflow specifications is provided by [10]. They use feature models [21] that are often used in software-product lines. The authors characterize their approach with the following words: “In fact, FCWs (Feature Configuration Workflows, as we propose them also provide a means to capture configuration knowledge and to make it accessible to non-experts. In addition, FCWs capture information about the variability of the system.” The authors use workflows to configure feature models. However, the opposite would be possible as well. Workflow specifications could be configured by feature models.

Feature models might be used in combination with generic components to manage different instances. They provide alternative configuration mechanisms and can be considered to be orthogonal to the discussed approach. It would be interesting to study whether feature models could be an accepted notation for adaptation of workflow specifications. However, this was out of the scope of the investigations of the current paper.

4 Summary and Outlook

In this paper we presented the idea of introducing generic components to BPMN based on the analysis that modeling of business processes is very time consuming and reuse of models is still difficult. The suggested approach is based on the concepts of design patterns, workflow patterns, task patterns and generic components.

An example from an OMG document was used to show how abstraction and generalization can help to get generic models that can be adapted to the necessary context. Adaptation of models can be described more formal, can be better controlled and is less error prone.

Even that the approach makes sense for working on paper tool support would be very helpful. It would allow seeing the instances of components with updated parameter values. User can see and evaluate their adapted current models.

Additionally, the concept of design-time and runtime parameter substitution was discussed.

Four strategies were identified.

1) Parameter substitution during design time
2) Parameter substitution during runtime
    d) Parameter substitution at start time fixed by the designer
    e) Parameter substitution at execution time fixed by the designer
    f) Parameter substitution at that time where parameters are available

A corresponding extension of BPMN by annotations was suggested.

The discussed generic components allow very flexible specifications and provide mechanisms that make specifications easier to be reused and the resulting specifications better readable.

References

1. BPMN, `http://www.bpmn.org` (last visited June 29, 2014)
2. BPMN 2.0 Tutorial: Get Started with Process Modelling using BPMN, `http://camunda.org/bpmn/tutorial.html` (last visited June 29, 2014)
3. OMG. BPMN 2.0 by Example, Version 1.0, OMG Document Number: dtc/2010-06-02, `http://www.omg.org/spec/BPMN/2.0/examples/PDF` (last visited March 5, 2014)
4. Breedvelt-Schouten, I.M., Paternò, F., Severijns, C.: Reusable structures in task models. In: Proc. DSV-IS 1997, pp. 225–239 (1997)
5. Brüning, J., Dittmar, A., Forbrig, P., Reichart, D.: Getting SW Engineers on Board: Task Modelling with Activity Diagrams. In: Gulliksen, J., Harning, M.B., van der Veer, G.C., Wesson, J. (eds.) EIS 2007. LNCS, vol. 4940, pp. 175–192. Springer, Heidelberg (2008)
6. CTTE, `http://giove.isti.cnr.it/tools/CTTE/last` (visited April 22, 2014)
7. P. Forbrig, C. Martinie, P. Palanque, M. Winckler: Rapid Task-Models Development using Sub-models, Sub-routines and Generic Components. submitted to HCSE 2014 in Paderborn (2014)

8. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design patterns: Elements of reusable object-oriented software. Addison Wesley (1994)
9. HAMSTERS, `http://www.irit.fr/ICS/hamsters/` (last visited April 22, 2014)
10. Hubaux, A., Classen, A., Heymans, P.: Formal Modelling of Feature Configuration Workflows. In: SPLC 2009 Proceedings of the 13th International Software Product Line Conference, pp. 221–230 (2009)
11. Laue, R., Awad, A.: Visualization of Business Process Modeling Anti Patters. Electronic Communications of the EASST X, 1/12–12/12 (2009)
12. Radeke, F., Forbrig, P.: Patterns in Task-Based Modeling of User Interfaces. In: Winckler, M., Johnson, H. (eds.) TAMODIA 2007. LNCS, vol. 4849, pp. 184–197. Springer, Heidelberg (2007)
13. Rational Rose, `http://www-03.ibm.com/software/products/en/rosemod` (last visited April 22, 2014)
14. Recker, J., Rosemann, M., van der Aalst, W., Mendlin, J.: On the Syntax of Reference Model Configuration – Transforming the C-EPC into Lawful EPC Models. In: Proc. of the Workshop on Business Process Reference Models (BPRM 2005), pp. 60–75 (2005)
15. Russell, N., ter Hofstede, A.H.M., van der Aalst, W.M.P., Mulyar, N.: Workflow Control-Flow Patterns: A Revised View, BPM Center Report BPM-06-22, BPMcenter.org (2006)
16. Russell, N., ter Hofstede, A.H.M., van der Aalst, W.M.P., Wohed, P.: On the Suitability of UML 2.0 Activity Diagrams for Business Process Modelling. In: Stumptner, M., Hartmann, S., Kiyoki, Y. (eds.) Proceedings of the Third Asia-Pacific Conference on Conceptual Modelling (APCCM 2006). CRPIT, vol. 53, pp. 95–104. ACS, Hobart (2006)
17. Scheer, A.-W.: Business Process Engineering: Reference Models for Industrial Enterprises. Springer (1994) ISBN-10: 3540582347
18. Scheer, A.-W., Jost, W., Güngöz, Ö.: A Reference model for industrial enterprises. In: Fettke, P., Loos, P. (eds.) Reference Modeling for Business Systems Analysis, pp. 166–181. Idea Group Publishing, Hershey (2007)
19. Scheer, A.-W., Feld, T., Caspers, R.: BPM: new architecture driven by Business Process Planning and Control (BPPC). IM Journal for Information Management and Consulting (2) (2012), special print, `https://www.scheerprocesstailor.com/tpl/download/IM-2012-Ausgabe2-Sonderdruck_EN_Homepage.pdf`
20. Scheer-Process-Tailor, `https://www.scheerprocesstailor.com/` (last visited: April 22, 2014)
21. Schobbens, P.-Y., Heymans, P., Trigaux, J.-C., Bontemps, Y.: Feature Diagrams: A Survey and A Formal Semantics. In: RE 2006, pp. 139–148 (September 2006)
22. Sinnig, D., Gaffar, A., Reichart, D., Seffah, A., Forbrig, P.: Patterns in Model-Based Engineering. In: Proc. CADUI 2004, pp. 195–208 (2004)
23. Together Software Modelling Tool, `http://www.step-10.com/SoftwareTools/Together/introduction.html` (last visited April 22, 2014)
24. UML, `http://www.uml.org` (last visited June 29, 2014)
25. van der Aalst, W.M.P., Basten, T.: Inheritance of Workflows: An Approach to Tackling Problems Related to Change. Theoretical Computer Science 270(1-2), 125–203 (2002)
26. van der Aalst, W.M.P., ter Hofstede, A.H.M., Kiepuszewski, B., Barros, A.P.: Workflow Patterns. Distributed and Parallel Databases 14(3), 5–51 (2003)

27. van der Aalst, W.M.P., Dreiling, A., Gottschalk, F., Rosemann, M., Jansen-Vullers, M.H.: Configurable Process Models as a Basis for Reference Modeling. In: Proc. of the Workshop on Business Process Reference Models (BPRM 2005), pp. 76–82 (2005)
28. White, S.A.: Process Modeling Notations and Workflow Patterns, BPTrends (March 2004), http://www.omg.org/bp-corner/bp-files/Process_Modeling_Notations.pdf (last visited at April 22, 2014)
29. Wohed, P., van der Aalst, W.M.P., Dumas, M., ter Hofstede, A.H.M., Russell, N.: On the Suitability of BPMN for Business Process Modelling. In: Dustdar, S., Fiadeiro, J.L., Sheth, A.P. (eds.) BPM 2006. LNCS, vol. 4102, pp. 161–176. Springer, Heidelberg (2006)

Contract Design and Uncertainty in Software Development Projects

Cornelia Gaebert

Research Group on Strategic Information Management,
European Research Center for Information Systems
c/o University of Muenster, Leonardo Campus 11, D-48149 Muenster, Germany
Cornelia.Gaebert@uni-muenster.de, Cornelia.Gaebert@indal.de

Abstract. Recent research has described software development projects in terms of the economic principal agent theory. These models mostly describe the customer as the principal, whereas the supplier is the agent. Our study shows, that regarding gaps in software requirement specifications, the supplier is in a principal situation, and the customer plays the role of an agent. Specifications are incomplete due to systematical reasons. Therefore, the customer must work on closing the gaps during the design and development phase of the project. From this, behavioral uncertainties arise. An empirical study supports our theoretical argument. We discuss consequences from these findings and derive suggestions for practitioners in software development projects.

Keywords: Incomplete Information, Principal-Agent Approach, Credible Contract, Software Development Project, Failure Reasons.

1 Introduction

Research on the contractual structure of software development projects, namely on outsourcing projects, has shown, that the relationship between the customer and the supplier can be described in terms of the economic principal agent theory [1, 2, 3]. The aim of these works is to use the findings of the economic theory for describing the contract structure of a software project in terms of the theory by understanding the behavior of the customer and the supplier as rational actors, and to make predictions for risk and success of a project with respect to the contractual situation.

Until now, most researchers describe the customer as the principal, and the supplier as the agent. This seems to be obvious, because the supplier works on behalf of the customer, having the needed information.

The aim of this paper is to show, that in a software development project there is also the contrary situation, where the customer is an agent, and the supplier is a principal. The root cause is that in the moment of signing the contract, the requirement specification is nearly never complete, and for systematical reasons, it cannot be complete. In this paper, we call this deficit requirement gaps.

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 217–230, 2014.

We will develop our argument in section 2 of this paper. In section 3, we will show some empirical findings, supporting our argument. Finally, in section 4, we will derive some consequences and conclusions and we will give a brief outlook for further research.

2 Principal and Agent in Software Development Projects

Researchers in the field of software project management focus mostly on the control of decisions and activities of the acting participants and stakeholders within the development organization [4, 5]. They often describe them as rational agents having goals and making decisions for the cooperation with other actors, with the purpose of achieving a maximum of benefit [6, 7]. However, as shown by Tollefsen [8], we also can consider organizations like companies or public authorities as rational agents having their own goals and making rational decisions for reaching these goals.

There is wide variety of cooperation structures between organizations in software projects. One software company may produce and deliver a system for many companies having the same requirements. In contrast, one company may engage more than one supplier in one project, or one supplier may engage some other companies or freelancers with a special knowledge within a project. Furthermore, a project may be processed completely within one company with its own development department for the implementation of information systems for several business departments. However, many software development projects are two-party projects like in the case of outsourcing [9]. For simplicity, the focus of this paper is on this kind of projects.

Regarding these software development projects, at the organizational level we can define two kinds of actors: First, there are organizations acting as customers; and second are the organizations acting as suppliers. The customer has business goals resulting in requirements for a software system, described in a requirement specification document. The supplier has the ability to develop an information system that meets these requirements. Therefore, the customer and the supplier sign a contract to carry out a software development project.

The supplier and the customer mostly agree on fixed-price contracts or contracts with a price ceiling (upper limit) [2, 10]. In these cases, the income of the supplier is nearly independent of its actual effort. Furthermore, if during the project new information regarding the requirements arise, the supplier will have additional expenses. In this paper we focus on projects under fixed-price conditions.

At first glance it seems to be obvious, that the supplier acts as an agent on behalf of the customer. The supplier has all the information needed for the development of the software system desired by the customer. The behavior of the supplier is not fully observable during the development phase, therefore, the customer is in a role called the principal by the economic principal agent theory. In recent years, researchers have shown that some parts of the structure of a software development project can be described by means of the principal agent theory, interpreting the customer as the principal, and the supplier as the agent.

Elitzur et al. [11] describe for outsourcing projects a double sided moral hazard problem, where also the customer is in an agent situation and the supplier works as a principal. The precondition stated by these authors is, that the outsourcing market is dominated by a few outsourcing companies. Therefore, these suppliers are in a strong position. As our empirical study shows (see section 3), in most cases of software development projects the customer is in a dominant role. We focus our work on this kind of projects.

In this section, we will develop the argument that in nearly all software development projects also the customer acts as an agent, and the supplier acts as a principal. To put it briefly, due to the fact, that requirement specifications are incomplete in most cases, the customer must work on closing gaps within the specifications. For the supplier, the efforts of the customer are not completely observable.

2.1 The Necessity of Incompleteness of Requirement Specifications

In an ideal world, the requirement specification is complete, unambiguous, and clear. In such a perfect world, the supplier has calculated all efforts for the implementation of the requirements before signing the contract. Based on the specification, the designers and developers will implement the needed system. No communication and no interaction between the parties will be necessary during the project.

Unfortunately, requirements are not complete and unambiguous. As shown in research literature [12, 13], and as stated by all experts in our empirical survey (see section 3), there are gaps in the requirements specifications. These gaps are one of the main reason for project failure. The Standish Group in 1995 [14] cited “incomplete requirements” as the number-one reason for failure, El Emam and Koru in 2008 [15] cited “too many requirements and scope changes.” Researchers and practitioners have exerted a lot of effort in developing methods for producing better specifications without gaps, misunderstandings, and unclear descriptions.

Nevertheless, there are systematic causes for the gaps in requirement specifications. Software requirement specifications contain knowledge in a strict sense only about the past and the present. For instance, the customer knows problems that exist with the currently used system, the present market situation, and business cases. About the future, there are only assumptions. In particular, how the new system will change the business processes is not a matter of fact, but a matter of expectation and anticipation. During the technical design process and the impementation phase of the project, questions regarding these assuptions arise, and the parties must decide on details not specified within the requirement specification document. In addition, the requirement engineer can only document consciously available knowledge, and to some extent subconsciously available knowledge. However, in all business processes, relevant conditions and information exist that no one knows about [16].

The customer has knowledge primarily regarding the business for which the software system is needed. In contrast, the supplier has knowledge regarding technical issues, like the properties of used frameworks and development techniques. Furthermore, on the supplier’s side, experiences from other projects regarding user acceptance and performance problems exist. This knowledge is also relevant for the development of a software system, but in the moment of documenting the requirements it is not avaialable.

Consequently, we have to accept the fact, that requirement specifications will contain gaps also in the future, and even if research in requirement engineering finds new and better methods, it will be a very long way to have complete requirement specifications in practice.

2.2 The Principal Agent Situation Regarding Incomplete Specifications

The choice of contract design depends on the expected uncertainty. Uncertainty occurs when customer and supplier close gaps in the requirement specification. The closing of gaps itself already represents an expense that supplier and the consumer have to pay during the course of the project. This task requires efforts from both sides in order to achieve an optimal result. The supplier has to deliver input such as technical information and the customer shall give business process details and professional information. Thus, asymmetrically distributed information [17] requires the closing of gaps in the specification during the development project. Interaction-related quality- and accountability problems are the consequence; interaction-induced behavioral uncertainty grows up [18]. Evidently, there is a great potential for conflicts.

Behavioral uncertainty is obviously, on both sides. Customers and suppliers have to quantify objectively all their supplies and efforts in order to measure performances and thereby reduce uncertainty [19]. Costs arise from procurement and measurement of information. These cost results from measurement itself, from loss through measurement errors or inaccuracies as well as exceeded costs through mutual opportunistically benefits from measurement errors and inaccuracies [20]. Therefore, prohibitively expensive measurement costs avoid the elimination of uncertainty. The contract has to deliver a compensation system that serves the dual function of allocating risks and rewarding cooperation.

The principal-agent approach is a central theoretical approach inside the new institutional economics. Here, the agent has an information advantage over the principal. Therefore, the principal has to manage his uncertainty about agent's behavior [21]. The agent knows earlier and more precisely his own intention and his own weaknesses. He uses this for his own advantage. The principal-agent approach concerns itself with issues and consequences from this. By contract design ex-ante provided penalties and incentive schemes are supposed to reduce uncertainty and their consequences during the execution of the contract. Thus, the contract reduces the risk of exploitation for the worse informed principal. In the best case, the contract protects against a misallocated contractual partner, economical disadvantages, and welfare loss.

Spremann [17] distinguishes between three basic types of behavioral uncertainty resulting from information asymmetries. These are (1) uncertainties about agent's capabilities to be able to perform the promised service (hidden characteristics), (2) uncertainties about agent's fairness to be willing to perform the promised service (hidden intention) and (3) uncertainties about agent's effort to perform the promised service (hidden action).

The terms revelation, authority, and incentive describe the possible designs of institutional hedging uncertainty. Revelation protects the principal against uncertainties about agent's capabilities (1). Authority protects the principal against uncertainties about agent's fairness (2). Incentive protects the principal against uncertainties about agent's effort (3).

For closing the gaps within the requirement specification, during system design and development the customer must answer questions, deliver business information, evaluate the suggestions of the supplier, and make decisions regarding software design issues. For the supplier, the processes within the organization of the customer are at least partly hidden. The supplier requests an information, and receives an answer, maybe as a document written by experts not known within the project. Obviously, the supplier is in a principal situation, whereas the customer is the agent working on behalf of the supplier. We will discuss now the cases known from the principal agent theory regarding behavioral uncertainty.

(1) *Hidden characteristics*. The supplier does not know if the customer is able to deliver the required information. Maybe the business information needed is not available from the business experts, or they have not the ability to understand the problem raising during the system design. Furthermore, it is possible, that the management of the customer is not able to organize the information generation and delivery in parallel to the everyday business.

For recognizing problems resulting from hidden characteristics, the principal agent theory describes signaling and screening. From this, we suggest for contract negotiation, that the customer shall state clearly the qualification of the experts available during the project for answering questions of the supplier (signaling). Furthermore, the customer shall describe the procedures of acquiring the needed information and shall commit on amounts of work time available for clarification. Consequently, the supplier is able to calculate the risk left from hidden characteristics. On the other side, the supplier should implement methods for getting informations regarding the relevant qualifications on the customer side (screening).

The contract design facing the problems from hidden characteristics is *revelation*. For the software development project, this means, that the parties define within the contract clearly the obligations of the customer for delivering information and that the customer commits the availability of experts with the needed qualification.

(2) *Hidden intention*. Because incomplete and ambiguous software specifications are subject to interpretation, the customer has the option to reject project results, stating that these results do not fulfill the requirements according to the interpretation of the customer. In such cases, the customer has no interest to make the specifications clear. He will use gaps in specifications for saving costs, arguing that the supplier does not have implemented the system in a way he should. For hidden intentions of this kind, the customer may remain undefined namely non-functional requirements like performance and usability requirements. By testing the system, the customer may argue that the performance of the system does not meet the needs of the business, and that the usability is not state of the art. With this argument, he can try to reduce the

price or to refuse the acceptance if he does not need the system anymore due to changes in the customer's business.

The principal agent theory suggests avoiding specific investments for reducing loss of investments on the principal's side. Obviously, for the supplier developing an individual software system it is nearly not possible to avoid specific investments. He cannot expect to be able to use the developed software within another project. Furthermore, the contract design suggested by the theory for handling hidden intention is authority. However, we can interpret the fixed price contract as an authoritarian contract, but with the customer as the authority, not the supplier. We must note that the case of hidden intention on the customer's side is highly risky for the supplier.

(3) *Hidden action*. If the contribution of the customer for filling requirement gaps and clarify open issues is not visible to the supplier, the latter cannot evaluate the real effort. Maybe, the experts do not deeply analyze the raising issues or do not search for relevant business facts but state own assumptions as objective facts. Furthermore, it is possible, that in the backstage of the project some stakeholders play their own game, trying to hinder the project's success.

According to the theory, the principal shall monitor the behavior of the agent to recognize hidden action. The best way to do this is to make the work of clarification of requirement issues in common teams. If this is not possible, we suggest the supplier to ask for detailed information regarding the question who was involved within the clarification process and to what extent.

The suggested contract design for preventing issues raising from hidden action is incentive. Nevertheless, it seems to be difficult for the supplier to agree incentive for the delivery of high quality specification clarifications. Furthermore, the effect of incentives for avoiding hidden actions on customer's side in a software development project is uncertain.

3 Empirical Support of the Theoretical Argument

We support our theoretical findings with an empirical survey. First, it is essential that the supplier gets a fixed price for realizing the software system, or that there is a cap on the effort-based price. If the customer would pay an effort-based price for all of the work done by the supplier, he could invest a lot of work needed for closing the gaps in requirement specifications. Furthermore, if the resulting software system would not meet the real requirements of the customer, the supplier could do any rework necessary for implementing the needed functionality. Second, are there gaps in the requirement specification delivered by the customer by signing the contract? Third, are there uncertainties on both sides, especially on supplier's side?

For this empirical part of our study, we conducted a two-step evaluation. First, we developed a questionnaire in the form of a standardized online survey as a special kind of standardized survey [22]. Next, we conducted personal interviews to deepen our understanding of the results from the questionnaire. The period of the evaluation was one year.

For the questionnaire, we chose the standardized online survey to give the respondents an opportunity to reflect and to question their own companies [23]. The format of the online survey itself was legitimate because the interviewees were an IT-savvy group. Open answers supplemented the closed questions to not be too restrictive and to gather the covered information [24]. In the following, we will analyze and interpret the results descriptively.

Experienced project participants on both sides (customer and supplier) were interviewed. The questionnaire had to take the management perspective into account. Because it is not possible to address trivially the population of all manufacturers and customers of custom software, and because questioning the population about any associated unacceptably high cost is not realistic, we chose a smaller population. Therefore, we could not achieve complete representativeness [23]. For practical reasons, we addressed the 45 members of a network of IT companies in Germany. Fifty additional addressees were available from other contacts. To expand the circle of respondents and to amplify the customer side, we used contacts in social networks such as Facebook (approximately 30), Xing (approximately 20), and Twitter (approximately 50). This ensured that the respondents had experience in different contexts of possible projects. Of the 200 addressees who were requested to participate in the survey, 29 actually completed the questionnaire (14 suppliers, 5 customers, 9 suppliers and customers (both), and 1 other).

A total of 48.3% of the respondents indicated that they belong to management and that they have responsibility for the contracts; 27.6% are project managers; 6.9% are employees at the working level; and 17.2% perform other activities, such as consulting. A total of 89.7% of the respondents had 10 or more years of experience with software development projects. The participants represented a broad range of sizes of projects with regard to the duration and number of employees.

For the exemplary and in-depth interviews, we conducted semi-structured expert interviews. We questioned, on the one side, a consultant with experience in software projects for approximately 15 years. He supports big companies in defining and organizing the contractual issues of software projects. On the other side, we spoke with a supplier with experience in software projects for approximately 20 years. He is an owner of a software development company with 10 programmers. Considering the sensitivity of failure research and the resulting difficulty in gaining access to project details, this methodology was most appropriate. The incomplete script of the semi-structured interview format left room for improvising questions [25]. The first interview lasted approximately 3 hours; the second lasted 1.5 hours. We made extensive notes during the interviews, which we evaluated afterward through a qualitative content analysis. Because we demanded appointed circumstances and facts, we avoided free interpretation problems [26].

3.1 Results from the Online Survey

The survey showed that the proportion of fixed-price contracts for software development projects is extremely high (Fig. 1). Taking into account that even the so-called agile fixed price, and time and material (T & M) price with ceiling ultimately

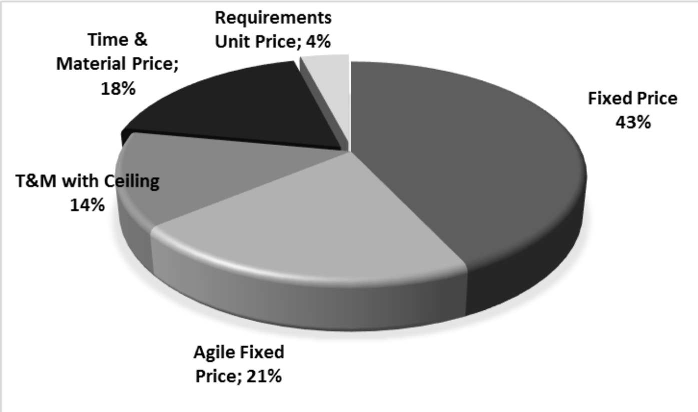


Fig. 1. Proportion of different types of contracts on software development projects

determines the maximum total budget for the consumer, the proportion of this type of contract is a total of more than three quarters of the software development projects. A manager on the side of the supplier added in free text: “Even if it is charged at T & M, the expectation of the customer is the compliance with the budget / value of the order.”

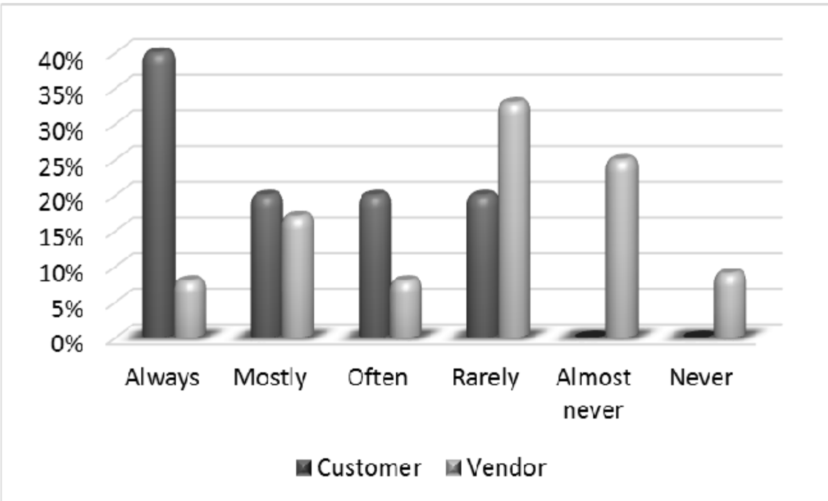


Fig. 2. Answer to the question "Do you determine the contract model?"

On the bottom line, the T & M price with ceiling and the agile fixed price mean the implementation of the requirements at fixed cost. Often the ceiling does not differ significantly from the calculated expense. An agile fixed price, however, allows one to the implementation of requirements when new requirements emerge. Then, these new requirements can replace earlier ones. However, such contractual subtleties relate only to new requirements. A third party (judge) can evaluate them. Nevertheless, this rarely helps in cases of closing the requirement gaps. Rather, closing gaps only makes unconscious knowledge aware. For the customer, it appeared typically obvious, whereas it was unknown to the supplier and vice versa. Filling the gaps makes it known explicitly.

The customers predominantly determine the contract model (Fig. 2). Although 80% of the customers indicate that they at least often determine the contract model, suppliers say quite the opposite. Two-thirds of them admit that they have little or no influence on the contract model. One comment from a project leader on the supplier side is: "I do not understand the question. The contract model is in all cases defined by the customer." Thus, customers clearly choose the contract design.

Customers and suppliers have different views on emerging problems inside a fixed-price project, like when an imbalance occurs in terms of time, cost, and quality (Fig. 3).

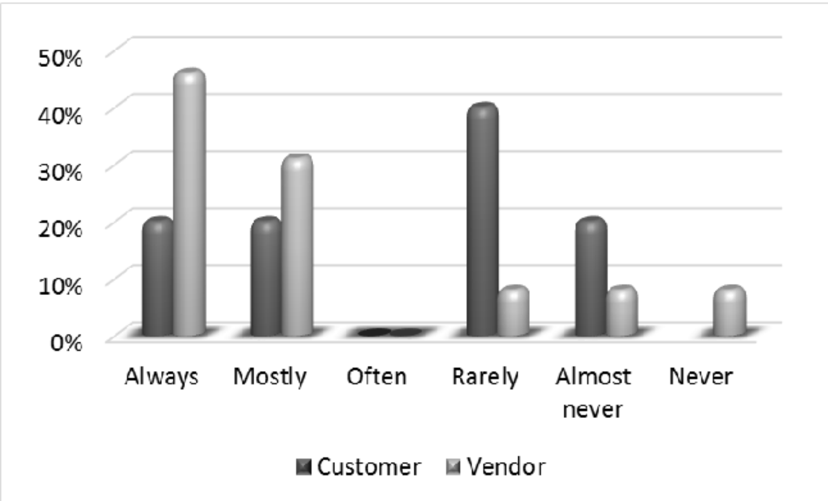


Fig. 3. Is an imbalance of time, cost, or quality in the project under fixed-price problematic?

Although 77% of the suppliers consider such a situation always or usually as problematic, 60% of the customers believe that this is rarely or almost never a problem for them.

Against this background, it is important to consider how the contract reflects gaps in the requirement specifications and how the signed contract supports the project itself. After all, such gaps lead to increased interaction. Most respondents stated for the vast number of projects (Fig. 4) that such gaps exist.

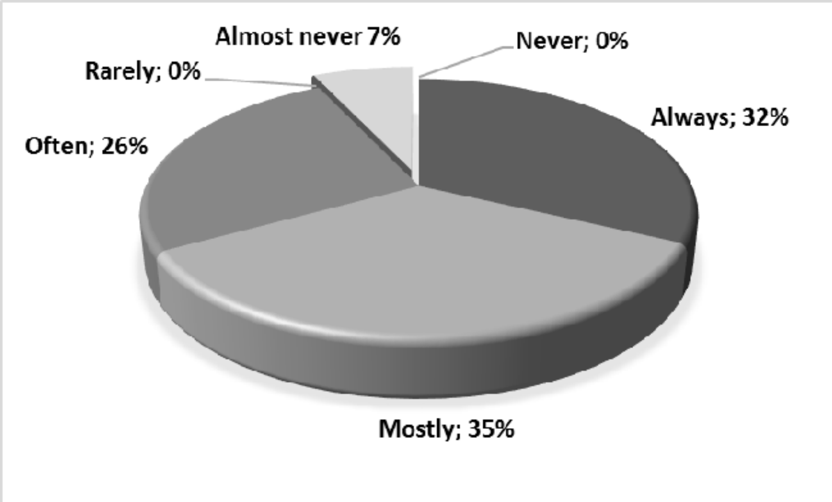


Fig. 4. Frequency of requirement gaps

Almost a third of the respondents said that such gaps “always” happen; 93% say that this case occurs at least often. However, a fixed-price contract hardly takes this sufficiently into account. For suppliers to do this seems hardly to be possible, as the notes to the relevant questions show. They try to work with a kind of overhead calculation but requirement gaps "are rarely sufficiently taken into account."

However, contracts widely do not reflect this fact. On the question, whether contractors continuously update the contract during the project, 81% of participants responded that this rarely or never happens.

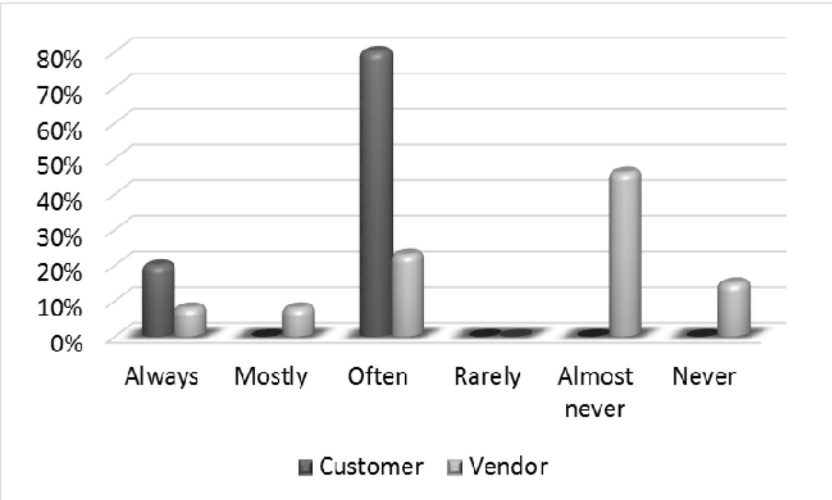


Fig. 5. Renegotiate customer and supplier requirement gaps

Customers and suppliers have a different perspective regarding whether demand gaps leading to unforeseen interaction would be renegotiated (Fig. 5). Although customers are of the opinion that this would always or at least often happen, 61% of the suppliers believe that there are never or almost never renegotiations.

Two-thirds of all respondents say that gaps in the requirement specifications always or almost always lead to unplanned discussions. The contract usually does not take into account the extra costs, which interactions trigger.

3.2 Results from Interviews

We documented the interviews in a structured way with references to each question and to the paragraph of the answer. In the following, we give a short overview of the results. In brackets, we note the reference to the minutes of the interviews. For example, (S Q3A2) references the supplier interview, question 3, answer paragraph 2.

Both interview partners said that the mostly preferred contract model is the fixed-price contract, especially if the requirements are documented and if they seem to be clear (C Q3A1, S Q5A2). This is because of the customer's restriction in having a limited budget and that customers must calculate the expected benefits against the costs beforehand (C Q16A1, (S Q6A1). Nevertheless, because "it is very seldom that the requirements are specified in a formal way" (C Q10A1), it is almost impossible to calculate the real costs. In addition, the supplier stated: "The problem does not come from the fixed price itself, but from unclear, incomplete, or changing requirements. And the problem is that the customer is not willing to change the price if he changes the requirements" (S Q6A3).

The interviews supported the finding from the online survey, that the customers mostly dominate the contract design(S Q5A2, C Q3A1). Nevertheless, both interview partners gave hints, that obligations for a cooperating behavior of the customer are possible in practice (C Q14A4, S Q11A6).

Because the requirement specifications were so important, we asked our interview partners to explain the reasons for the gaps, the possibilities for dealing with these gaps, as well as the consequences. Both sides cited the reasons as being "special" or "exceptional use cases" that the experts were not aware of during the requirements analysis or were too difficult to model (C Q11A1; C Q11A4; S Q10A1). Furthermore, the facts were "obvious" (C Q11A3) or "self-evident" (S Q10A1) to the business experts, so they did not speak about them. Nonfunctional requirements were often unknown to the users (S Q10A1).

Both interview partners showed a high degree of uncertainty regarding the behavior, intentions, and skills of the other side. Customers try to get certainty beforehand from information like "descriptions of credential projects, facts about the know-how of their staff, information about the methods in designing and processing a software project" (C Q7A1). With "governance structures for the project" (C Q5A1) the customer hopes to "get at early phases of the project a good feeling of the progress and the quality of the vendor's work" (C Q6A1). However, uncertainty remains high: "Nearly nobody can distinguish the clever, good one from the slow and poor one. And if the vendor mentions that there are unforeseeable problems, you don't know if he is right or he is not professional enough for doing his job" (C Q16A1).

Regarding the same issue, the interview partner from the supplier side said, "a new management, problems in his market, new relevant law, and maybe, the customer does not need the software anymore or the costs will be higher than the effects. Then, maybe, the customer's management tries to cancel the project" (S Q11A5).

On the customer side, the strategy is to handle all problems in a formal way and to avoid all discussions regarding efforts in narrowing the gaps in the requirements (C Q11A5; C Q15A1). In contrast, the supplier obviously has strategies of its own, knowing that the customer cannot see all that the supplier is doing (S Q12A1).

4 Conclusions and Further Research

As our empirical study shows, most contracts for software development projects are fixed price projects, mostly with a predefined price for the information system to be implemented, sometimes with a time & material price with ceiling or with an agile fixed price. Furthermore, the customer is in a dominant, authoritarian position, defining the contract design and the conditions of cooperation. If problems arise during the project, the consequences for the supplier can be hard, whereas the customer is often in a better position.

On the other hand, the empirical investigation shows that there are nearly in every project gaps in the requirement specifications. The customer must take part in closing these gaps during the development of the project. Therefore, the supplier is in a principal situation regarding the activities for closing requirement gaps, whereas the customer is the agent regarding the task of providing the needed information.

As we have shown during our theoretical argumentation, the supplier has some possibilities for dealing with the situation being the principal and the uncertainties raising from the information asymmetry. Regarding the problem of *hidden characteristics*, we suggest to ask the customer for details of the qualifications of the experts before signing the contract. Signals for qualifications can be experiences from other projects or knowledge in software requirement engineering.

Regarding the possibility of *hidden action*, we suggest the supplier to participate in the work of filling the gaps by doing the investigation needed in common teams. In this way, the supplier can monitor the work of the customer. If such a participation is not possible, the supplier should ask for details regarding the process of collecting and providing the information: Who was involved? Where are the sources of information?

As our empirical study shows, the customer is the dominant party in most projects. Therefore, for the supplier it is hard to deal with problems arising from *hidden intentions*. On one side, the contract situation hinders the supplier to implement authority. On the other side, in an individual software development project the supplier the supplier has no chance to avoid specific investments, because the software system to be implemented shall meet the specific requirements of this customer, and therefore the chance for using the system within another project are small.

Further research therefore shall focus on the problem of possible hidden intentions of the customer in software development projects. Furthermore, empirical studies shall investigate, which tools of signaling, screening and monitoring are possible in software development projects. With results from such investigations, suggestions for the contract design and for the project management process can be derived. Practitioners can avoid the problems from the principal agent situation described in this study by using these tools.

References

1. Keil, P.: Principal agent theory and its application to analyze outsourcing of software development. ACM SIGSOFT Software Engineering Notes 30(4), 1–5 (2005)
2. Lichtenstein, Y.: Puzzles in Software Development Contracting. Communications of the ACM 47(2), 61–65 (2004)
3. Basu, V., Lederer, A.L.: Agency theory and consultant management in enterprise resource planning systems implementation. ACM SIGMIS Database 42(3), 10–33 (2011)
4. Keil, M., Smith, H.J., Pawlowski, S., Jin, L.: 'Why Didn't Somebody Tell Me?': Climate, Information Asymmetry, and Bad News About Troubled Projects. SIGMIS Database 35(2), 65–84 (2004)
5. Rustagi, S., King, W.R., Kirsch, L.J.: Predictors of Formal Control Usage in IT Outsourcing Partnerships. Information Systems Research 19(2), 126–143 (2008)
6. Yilmaz, M., O'Connor, R.V., Collins, J.: Improving Software Development Process through Economic Mechanism Design Communications. Computer and Information Science 99, 177–188 (2010)
7. Cockburn, A.: The End of Software Engineering and the Start of Economic-Cooperative Gaming. ComSIS 1(1) (2004)
8. Tollefsen, D.: Organizations as true believers. Journal of Social Philosophy 33(3), 395–410 (2002)
9. Chua, C.E.H., Lim, W.-K., Soh, C., Sia, S.K.: Client strategies in vendor transition: A threat balancing perspective. The Journal of Strategic Information Systems 21(1), 72–83 (2012)
10. Oestereich, B.: Der agile Festpreis und andere Preis- und Vertragsmodelle. Objekt-Spektrum 01, 29–33 (2006)
11. Elitzur, R., Gavious, A., Wensley, A.K.P.: Information systems outsourcing projects as a double moral hazard problem. Omega 40, 379–389 (2012)
12. Liu, J.Y.-C., Chen, H.-G., Chen, C.C., Sheu, T.S.: Relationships among interpersonal conflict, requirements uncertainty, and software project performance. International Journal of Project Management 29, 547–556 (2011)
13. McGee, S., Greer, D.: Towards an understanding of the causes and effects of software requirements change: two case studies. Requirement Engineering 17, 133–155 (2012)
14. Standish Group: CHAOS Report (1995), `http://www.projectsmart.co.uk/docs/chaos-report.pdf` (June 21, 2011)
15. El Emam, K., Koru, A.G.: A Replicated Survey of IT Software Project Failures. IEEE Software 25(5), 84–90 (2008)
16. Kano, N., Seraku, N., Takahashi, F., Tsuji, S.: Attractive Quality and Must Be Quality. Quality - Journal of the Japanese Society for Quality Control 14(2), 39–44 (1984)

17. Spremann, K.: Asymmetrische Information. ZfB 60(5/6), 561–586 (1990)
18. Kaas, K.P.: Informationsökonomik. In: Tietz, B., Köhler, R., Zentes, J. (eds.) Handwörterbuch des Marketing. 2., völlig neu gestaltete Auflage. Schäffer-Poeschel, Stuttgart (1995)
19. Barzel, Y.: Measurement Cost and the Organisation of Markets. Journal of Law & Economics 25(1), 27–48 (1982)
20. Alchian, A., Demsetz, H.: Production, Information Costs, and Economic Organization. American Economic Review 62(5), 777–795 (1972)
21. Arrow, K.J.: Agency and the Market. In: Arrow, K.J., Intriligator, M.D. (eds.) Handbook of Mathematical Economics, vol. 23(3). North-Holland Publ. Co., Amsterdam (1986)
22. Klammer, B.: Empirische Sozialforschung. Eine Einführung für Kommunikationswissenschaftler und Journalisten. Utb, Konstanz (2005)
23. Schnell, R., Hill, P., Esser, E.: Methoden der Sozialforschung. 9. Auflage. Oldenbourg Wissenschaftsverlag, München (2011)
24. Mayer, H.: Interview und schriftliche Befragung. Entwicklung, Durchführung und Auswertung. Oldenbourg Wissenschaftsverlag, München (2012)
25. Myers, M.D., Newman, M.: The qualitative interview in IS research: Examining the craft. Information and Organization 17(1), 2–26 (2007)
26. Gläser, J., Laudel, G.: Experteninterviews und qualitative Inhaltsanalyse. 4. Auflage. VS Verlag, Wiesbaden (2010)

Assessing Aspect Oriented Approaches in Business Process Management

Amin Jalali

Department of Computer and Systems Sciences,
Stockholm University
aj@dsv.su.se

Abstract. Separation of concerns is an important topic in Business Process Modeling. One sort of concerns is cross-cutting, like security, which are repeated in many business processes. These concerns make the models more complex, since concerns are repeated in many process models. The repetition of realization of concerns in process models makes the maintenance cumbersome. Aspect Oriented Business Process Modeling is an approach to address these concerns, which has been investigated recently. However, no set of requirements are defined for such modeling proposals, which makes the evaluation of and comparison between these approaches impossible. Therefore, this paper introduces a set of requirements for the aspect oriented business process modeling, which are used to define an evaluation framework for assessing these modeling approaches. The framework is used to evaluate existing aspect oriented business process modeling proposals. The result shows a comparison between different modeling proposals by clarifying their strengths and weaknesses. It also shows the gap in the area, which can be used as direction for future research.

Keywords: Business Process Modelling, Aspect Oriented, Requirements, Evaluation.

1 Introduction

Business Process Modeling is an important area, aims to enhance the (re-)design of business processes to be more efficient. Processes can be very complex, which makes the process models complex as well. This complexity hinders the comprehension of process models, so different techniques are used in process modeling to deal with the complexity like separation of concerns. By separating concerns, people can deal with less complex modules at a time, which enhance their capability to understand the process models. This ability empowers people to redesign processes to improve efficiency.

Process models contain different concerns, which can be separated through three medialization techniques such as vertical, horizontal and orthogonal [11]. *Vertical Modularization* aims to hide process details by introducing subprocesses, which improves the modeling structure of a process for dealing with

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 231–245, 2014.

complexity. The domain of a sub-process is limited to one process, and it can be re-used several times in the process. *Horizontal Modularization* aims to facilitate dealing with the complexity of process models by introducing peer modules. These modules can be considered as sub-processes, which are common for different process models and are not limited to one process. In this way, a peer module can be re-used in different processes. *Orthogonal Modularization* aims to separate the dependency of process models from peer modules, so the relation between peer modules and process models would be documented using some rules. In this way, the usage of peer modules can be altered by changing the rules, rather than updating all process models. These techniques can be used to separate different types of concerns.

Different techniques are introduced for modeling different sort of concerns, among which Aspect Oriented Business Process Modeling aims to separate cross-cutting concerns from process models. Cross-cutting concerns are those which their realization in process models results in scattering and tangling problem. *Scattering* means that the realization of concerns should be repeated in different process models, e.g. different process models should comply to one security concern. *Tangling* means that the changes in application of the concern need to be reflected in different process models, e.g. if the security concern does not need to be applied for the payment, all processes which have payment mechanism should be found and untangled. Aspect oriented business process modeling aims to separate both realizations of these concerns and their application, which improves re-usability, maintenance, and dealing with complexity [2,4,8].

Although different modeling approaches are proposed for aspect oriented business process modeling , there is not any comprehensive definition of requirements to specify the characteristics that these models should have. This gap hinders the evolution of this sort of modeling, since it is not possible to consider the missing parts in proposed approaches. Moreover, it is not possible to compare these approaches with each other to investigate which one has a better degree of support for separation of cross-cutting concerns.

Therefore, this paper defines a set of requirements for Aspect Oriented Business Process Modeling. It defines a framework based on these requirements that can be used for evaluating this kind of modeling approaches. It also evaluates existing aspect oriented business process modeling approaches and shows the pros and cons of each approach. The result of this evaluation shows the missing areas which require more research and further investigation.

Thus, the paper is organized as follows. Section 2 introduces the basic modeling structures which are needed for Aspect Oriented Modeling. Section 3 defines a set of requirements for aspect oriented business process modeling, which should be considered when defining a modeling notation. These requirements are used to construct a framework for evaluating the aspect oriented business process modeling approaches in section 4. In section 5, an evaluation of ten existing modeling approaches is given. Finally, section 6 concludes the paper and introduces the direction for future works.

2 Background

This section introduces basic constructs of Aspect Oriented Business Process Modeling through an example. These constructs are common in aspect oriented approaches in different disciplines such as programming [10], which are aimed to solve scattering and tangling problem in Business Process Management (BPM) by enable separation of cross-cutting concerns from process models. Fig. 1 represents a motivating example which is used in this section to facilitate depiction of the problem and its solution.

The example in Fig. 1 contains two sides: the left side of the figure shows the relation between cross-cutting concerns and some process models; the right side of the figure shows a fictitious process model, called `Transfer Money Process`. As it can be seen on the left-hand side, a concern can be repeated in many process models, which result in scattering problem. The process on the right side of the figure starts when a customer fills a form. If the customer wills to transfer money to his or her account, no security control is taken. Otherwise, the customer should sign the transaction, and the request is investigated for potential fraud detection. Afterwards, the money is transferred. As a part of security concern, the customer will be notified about the transaction if the destination account is owned by another person. Finally, the transaction should be logged. This process is tangled to logging and security concerns, which means that the activities which are needed to represent the process model is twisted with concerns' activities.

The scattering and tangling problems result in many difficulties in designing and maintaining business processes. For example, if the security aspect is changed in this example, all processes should be examined to consider if they should be also altered. This can result in inconsistency between processes and different aspects, if a realization of an aspect is forgotten to be updated in one process model. In this way, managing changes in cross-cutting concerns are cumbersome and hinders supporting modeling and enactment of business processes. The aspect orientation aims to solve the scattering and tangling problem.

To solve the scattering and tangling problems, both cross-cutting concerns and their dependencies to process models should be separated from process definition. Therefore, orthogonal medialization can be applied to solve this problem through encapsulating cross-cutting concerns and the dependency rules into

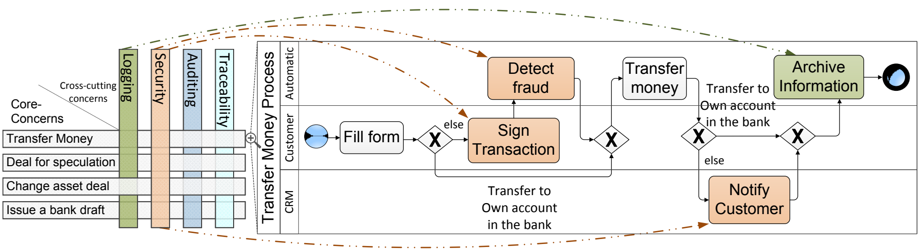


Fig. 1. Cross-cutting concerns in Transfer Money Process

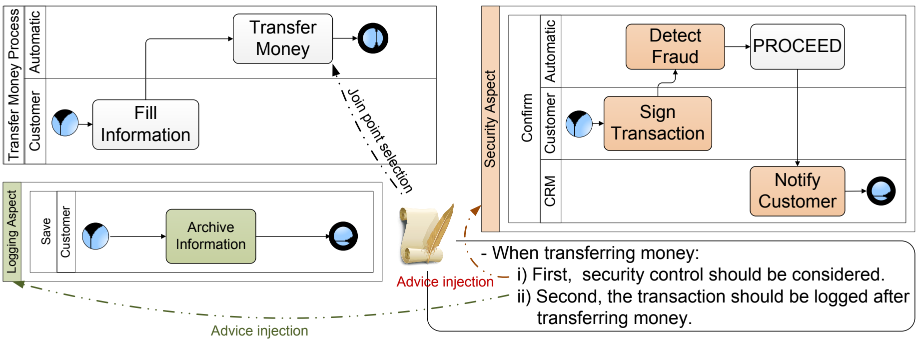


Fig. 2. Aspect Oriented Modeling in BPM

separated modules. In this way, cross-cutting concerns and their dependencies can be changed without hindering process model's definitions.

Cross-cutting concerns can be separated through encapsulating them into individual peer modules, called advices. An *advice* is a process model which has a start and an end event. It also contains activities, which must be performed for realizing a cross-cutting concern. Fig. 2 shows an aspect oriented model for the given motivating example, where `Confirm` and `Save` processes represent the realization of a security and logging cross-cutting concerns respectively. A set of advices with a common goal can be grouped together in a module called *aspect*. In this example, `Confirm` advice belongs to `Security Aspect` and `Save` advice belongs to `Logging Apect`.

The dependency between cross-cutting concerns and process models can be separated through definition of rules, called *Pointcuts*. Each pointcut specifies rules, which indicate points in process models for them an advice should be applied. The potential points in process models that a pointcut can select is called *join points*. These points are activities in a process model, e.g. `Fill Information` and `Transfer Money` activities in Fig. 2. The point which is selected by a pointcut is called *advised join point*, and the selection process is called *join point selection*. For example, `Transfer Money` activity is selected by the defined pointcut in Fig. 2. An advice can be considered before, after or around an advised join point [4]. To enable the around scenario, the place of the advised join point can be defined using a placeholder in the advice process, called `PROCEED` (see `Confirm` advice in Fig. 2).

Each pointcut can also define the relation between cross-cutting concerns for a specific point in a process model. Advices can have the same order or different orders [8]; for example, the security concern in Fig. 2 should be considered before the logging concern. This information can be specified in pointcut definition. Through this definition, pointcut specifies how different advices should be related to advised join points. This process is called *advice injection*.

There are different aspect oriented business process modeling approaches, which are defined based on this set of requirements. Although they mostly support definition of these elements, their capability in supporting separation of cross-cutting concerns are different. The difference is rooted in the set of requirements that they support. To evaluate their separation level and enable comparison between them, next section introduces sets of requirements for aspect oriented business process modeling.

3 Requirements

This section introduces a set of requirements that an aspect oriented business process modeling should fulfill to enable supporting cross-cutting concerns in BPM area. The requirements are compiled through literature review, and lesson learned from case studies by enabling aspect orientation [8,7,9]. Two sets of requirements are defined. The first set is compulsory without them aspect oriented business process modeling cannot be enabled in BPM. The second sets are optional, which enable assessing the level of separation of cross-cutting concerns from process model for each approach.

3.1 Basic Requirements (B.R.)

Four important perspectives are recognized in BPM area to specify a business process, i.e. functional, control-flow, data and resource perspectives. Functional perspective describes the activities that a process contains. Control-flow perspective specifies the order between these activities. Data perspective indicates the required information to perform an activity. Resource perspective describes the resource (person or system) that should perform the activity. An aspect oriented business process modeling aims to extend the functionality of process modeling to support separation of cross-cutting concerns. Thus, it should support these perspectives as well. Moreover, cross-cutting concerns can be separated if the solution addresses the scattering and tangling problem, so it is a basic requirement for every Aspect Oriented Business Process Modeling approach to address scattering and tangling problem. Therefore, three basic requirements can be defined for every Aspect oriented Business Process Modeling approach such as:

– B.R.1. Aspect Oriented Business Process Modeling should support definition of business processes using functional, control-flow, data and resource perspectives.
– B.R.2. Aspect Oriented Business Process Modeling should remove scattering problem in definition of concerns in process models.
– B.R.3. Aspect Oriented Business Process Modeling should remove tangling problem in definition of concerns in process models.

3.2 Measurement Requirements (M.R.)

The requirements for assessing the maturity of an aspect oriented business process modeling approach can be defined based on three issues: i) the strongness of join point selection, ii) the strongness of advice injection and iii) the available support for different phases in BPM lifecycle. The join point section depends on two sub-issues, i.e. *Signature Exposure* and *Rule Composition*. The advice injection depends on two sub-issues, i.e. *Pointcut Definitions* and *Transformation Patterns*. These issues and their requirements are explained below.

Signature Exposure (M.R.S.). The first step in join point selection is recognizing the points in process models to which an advice can be related. These points can be exposed by process models based on different business process perspective, i.e. functional, control-flow, data and resource. The points which can reveal information about each of these perspectives are called signatures. This term is borrowed from Aspect Oriented Programming paradigm [10]. There are four types of signature that can be defined in aspect oriented business process modeling:

- *M.R.S.1 Process: the approach should expose points of control-flow perspective of processes for which a concern can be defined.* The process name can be considered as an exposure point, which enables the definition of the relation between cross-cutting concerns to a process model.
- *M.R.S.2 Tasks: the approach should expose points for functional perspective for which a concern can be defined.* For example, the names of activities are candidates for task signatures.
- *M.R.S.3 Data: the approach should expose points for data perspective for which a concern can be defined.* For example, reading and writing a data entity can be defined as data signatures in process models.
- *M.R.S.4 Resource: the approach should expose points for resource perspective for which a concern can be defined.* For example, the resource name and role can be considered as resource signatures.

Rule Composition (M.R.R.). The join point selection is derived through interpreting pointcut rules when an instance of a process model is enacted, and the rules play an important role in enabling separation of cross-cutting concerns. The rules can be defined using information about different perspectives of a process model. The degree of separation is related to the number of perspectives that can be used in composing rules. Moreover, it is important if a pointcut can be defined based on a composition of different perspectives' information, and if there is a dominant perspective in composition of rules. A dominant perspective is a perspective that all other perspectives should be defined based on it. For example, functional perspective is considered as a dominant perspective in rule composition in AO4BPMN [4]. The existence of a dominant perspective result in definition of rule redundancy [7]. Therefore, the following requirements can be considered to evaluate the rule composition dimension.

- *M.R.R.1 Process: the approach should support definition of rules based on control-flow perspective information.*
- *M.R.R.2 Task: the approach should support definition of rules based on functional perspective information.*

- *M.R.R.3 Data: the approach should support definition of rules based on data perspective information.*
- *M.R.R.4 Resource: the approach should support definition of rules based on resource perspective information.*
- *M.R.R.5 Combination: the approach should support composition of rules based on combinations of different process perspectives.*
- *M.R.R.6 Domination: the approach should support composition of rules without any dominant perspective.* For example, it should be possible to define a rule based on resource perspective information without mentioning the task information.

Advice Relations (M.R.A.). The level of separation is also affected by the ability to define a relation between an advice and i) different join points, and ii) other advices. An advice can be defined to be considered before, after or around a point in a process model. Moreover, an advice can be defined in parallel with other advices, or it can be defined for another advice (nested scenario). It is also possible to have precedence between advices when they are related to a point in a process model.

- *M.R.A.1 Before: the approach should enable definition of before advices.* Before advices are those which are considered before a join point.
- *M.R.A.2 After: the approach should enable definition of after advices.* After advices are those which are considered after a join point.
- *M.R.A.3 Around: the approach should enable definition of around advices.* Around advices are those which are considered around a join point.
- *M.R.A.4 Parallel: the approach should enable definition of parallel advices for a join point.*
- *M.R.A.5 Nested: the approach should enable definition of nested advices.* Nested advices are those which are defined for another advice.
- *M.R.A.6 Precedence: the approach should enable definition of precedence between advices for a join point.*

Transformation Patterns (M.R.T.). In aspect oriented business process modeling, every concern is encapsulated into individual modules. Although this approach makes coping with the complexity easier, it needs transforming knowledge from one module to another. The knowledge can be related to different perspectives, which can be articulated as transformation's patterns. The more pattern supported by an approach means better support for separation of crosscutting concerns.

- *M.R.T.1 Process: the approach should enable transformation of process level data among different related modules.*
- *M.R.T.2 Task: the approach should enable synchronization of PROCEED placeholders in advices with advised join point.*
- *M.R.T.3 Data: the approach should enable transformation of data among different related modules.*

– *M.R.T.4 Resource: the approach should enable transformation of resources which has performed activities among different related modules.*

Phases Support (M.R.P.). It is important that this modularization technique can be supported in different BPM lifecycle such as design, run and adjustment. The adjustment should be performed for running process instances, and different sort of adjustments can be defined based on the fact that whether the core-functionalities of the business process and cross-cutting concerns are already started at the time of adjustment or not. Therefore, three kinds of adjustment can be defined (see Fig. 3):

– *Backward Adjustment* in which cross-cutting concerns should be adjusted when both the core-functionalities of process model and cross-cutting concerns are enacting. For example, changing the security concern when a purchase process and its related security concern are running.
– *Backward-Forward Adjustment* in which cross-cutting concerns should be adjusted when the core-functionalities of a process model is running, but the cross-cutting concerns realizations are not. This scenario happens when the advised join point in aspect oriented process model is not yet enabled. For example, changing a security concern when a purchase process is running, but its security concerns have not yet run.
– *Forward Adjustment* in which cross-cutting concerns should be adjusted when neither the core-functionality of process model nor cross-cutting concerns realizations are enacting. For example, changing a security concern when a purchase process is not yet started.

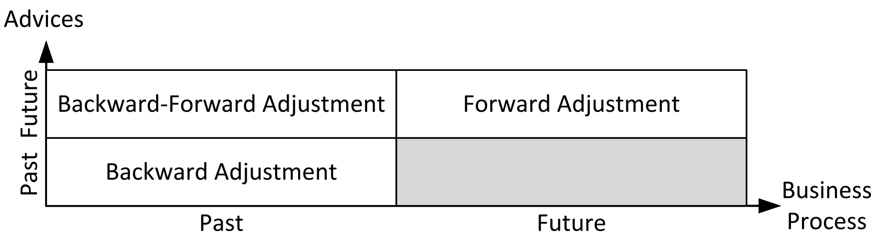


Fig. 3. Adjustment Types

It should be noted that it is not possible to have a scenario in which a cross-cutting concern is running while its business process has not yet started, so the related section is grayed out in the figure. Therefore, the Phases support requirements can be defined as below.

– *M.R.P.1 Design: the approach should support the design of aspect oriented business process modeling.*
– *M.R.P.2 Run: the approach should support enactment of aspect oriented business process models.*

- *M.R.P.3 Backward Adjustment: the approach should support adjustment of running advices.*
- *M.R.P.4 Backward-Forward Adjustment: the approach should support adjustment of new advices for running cases.*
- *M.R.P.5 Forward Adjustment: the approach should support adjustment of advices for new cases.*

4 Evaluation Framework

This section proposes a framework based on defined requirements. This framework can be used for evaluating Aspect Oriented Business Process Modeling approaches. The evaluation contains two steps as mentioned in previous section, i.e. selection and measurement.

In selection step, the approach is evaluated based on basic requirements (B.R.1-B.R.3), i.e. it should enable definition of business processes using control-flow, functional, data and resource perspectives, and it should also address the problem of scattering and tangling of cross-cutting concerns in process models.

To evaluate each approach based on defined measurement requirements,

- Let $M.R.m$ denotes a set of Measurement Requirements, where m is a variable referring to the sub-issues introduces in section 3.2.
  – It means that $M.R.m$ is a set of Measurement Requirements including $M.R.S.$ (Signature Exposure), $M.R.R.$ (Rule Composition), $M.R.A.$ (Advice Relations), $M.R.T.$ (Transformation Patterns) and $M.R.P.$ (Phases Support).
- Let $M.R.m.i$ denotes the ith requirement in $M.R.m$, where i is an integer and $1 \leq i \leq |M.R.m|$
- Let $E.R.m$ denotes Evaluation of Requirements, where m is a variable referring to the specific Measurement Requirement set, i.e. $M.R.m$.

The $E.R.m$ can be calculated using this formula:

$$E.R.m = \Big(\frac{(\sum_{i=1}^{n} M.R.m.i) \times 4}{|M.R.m|}\Big) \tag{1}$$

The result can be illustrated using a pentagon containing five dimensions each of which represents a set if of requirements (see Figure 4). The next section shows the evaluation of current approaches in aspect oriented business process modeling using this framework.

5 Evaluation Results

This section shows the evaluation result of applying the framework on different aspect oriented business process modeling approaches. The approaches are selected as a result of studying different literature about aspect orientation in BPM area. The evaluation of approaches contains two steps as mentioned in previous section, i.e. selection and measurement.

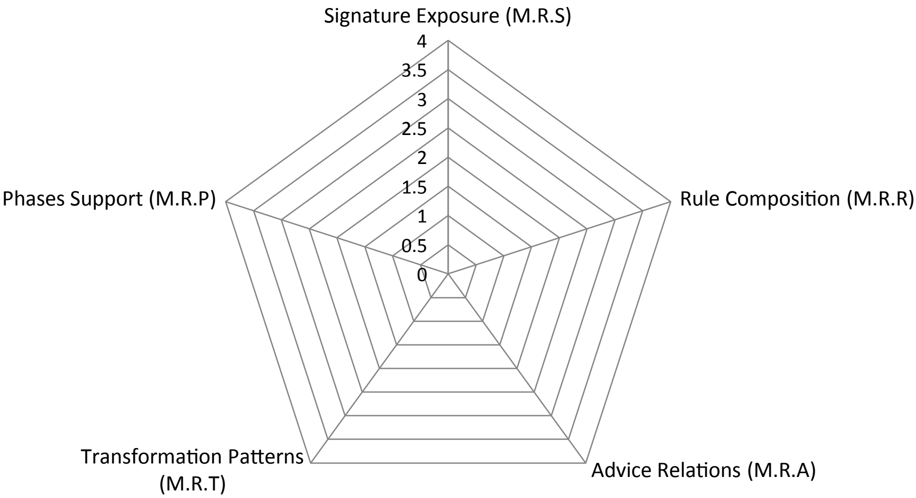


Fig. 4. The illustration for Evaluation Framework result

5.1 Selection Step Result

In selection step, ten aspect oriented business process modeling approaches are identified through studying literature. These approaches are evaluated based on basic requirements. The result is shown in Fig. 5. As it can be seen in the figure, half of these approaches do not meet the basic requirements, where:

- The approach proposed by Charfi et al. (AO4BPEL) [3] does not meet the B.R.1, i.e. it does not consider the support for resource perspective.
- The approach proposed by Wang et al. [15] neither meet B.R.1 nor B.R.3, i.e. it does not consider the support for resource perspective, and it relates the process models to cross-cutting concerns through some elements (called lose and gain), which introduces the tangling problem.
- The approach proposed by Shankardass [14], Jalali et al. (AOBPMN) [8] and Collell [5] do not meet the B.R.3 since they introduce elements in the main process model, which introduces the tangling problem.
  – Shankardass [14] introduces dot points in process models as a means to relate processes to cross-cutting concerns.
  – Jalali et al. (AOBPMN) [8] and Collell [5] introduce intermediate conditional events as a way to related process models to cross-cutting concerns.

The other five approaches meet the basic requirements, so they can be assessed using the proposed framework which is explained in the next section.

5.2 Measurement Step Result

In this section, each of five approaches which met the basic requirements are assessed. The results are explained below.

Approaches	Basic Requirements (B.R.)			Measurement Requirements (M.R.)											
				Signature Exposure (M.R.S.)					Rule Composition (M.R.R.)						
	1	2	3	1	2	3	4	result	1	2	3	4	5	6	result
Charfi et al. (AO4BPEL)		1	1												
Wang et al.		1													
Shankardass	1	1													
Collell	1	1													
Jalali et al. (AOBPMN)	1	1													
Charfi et al (AO4BPMN)	1	1	1		1			1		1					0.667
Jabeen et al.	1	1	1		1			1		1					0.667
Patiniotakis et al.	1	1	1		1			1		1					0.667
Cappelli et al. (AO-BPMN)	1	1	1		1			1		1	1		1	1	2.667
Jalali et al.	1	1	1		1			1		1	1		1		2

Approaches	Measurement Requirements (M.R.)																	
	Advice Relations (M.R.A.)							Transformation Patterns (M.R.T.)					Phases Support (M.R.P.)					
	1	2	3	4	5	6	result	1	2	3	4	result	1	2	3	4	5	result
Charfi et al. (AO4BPEL)																		
Wang et al.																		
Shankardass																		
Collell																		
Jalali et al. (AOBPMN)																		
Charfi et al (AO4BPMN)	1	1	1	1			2.667		1			1	1					0.8
Jabeen et al.	1	1	1	1			2.667		1			1	1					0.8
Patiniotakis et al.	1	1	1				2		1			1	1					0.8
Cappelli et al. (AO-BPMN)	1	1		1			2		1			1	1					0.8
Jalali et al.	1	1	1	1			2.667		1	1		2	1	1		1	1	3.2

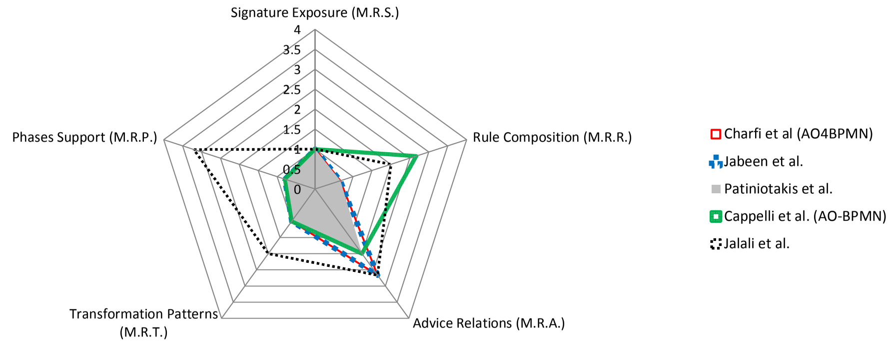


Fig. 5. EvaluationResult

1. Charfi et al. (AO4BPMN) [4] extended Business Process Model and Notation (BPMN) [12] to support separation of cross-cutting concerns from process models. This approach only considers the task signature when defining join points, so it only meets one of the requirements in the Signature Exposure (M.R.S.) set. In rule composition, it simply defines the composition of rules based on task names. Although the authors are aware of other elements, which can be considered in rule definition, they just realize the task information in the Rule Composition (M.R.R.) set for simplicity. AO4BPMN meets four requirements from the Advice Relations (M.R.A.) set, i.e. before, after, around relations between advices and process models. It also supports definition of parallel advices. However, it neither defines the nested relation nor the precedence between advices. This approach only defines how tasks between process models and advices are related to eachother, so it merely considers the task transformation pattern from the set of Transformation Patterns (M.R.T.). It also only supports the design phase of BPM lifecycle which meets just one of the requirements from the Phases Support (M.R.P.) set.
2. Jabeen et al. [6] propose an approach based on AO4BPMN in which a concrete language is defined for expressing pointcuts. However, they did not consider any additional requirements when composing rules, which makes the support level of separation of cross-cutting concerns for this approach equivalent to AO4BPMN. Thus, the graph of both this approach and AO4BPMN is identically the same, as it can be seen in Fig. 5.
3. Patiniotakis et al. [13] try to extend AO4BPMN with introduction of some new elements like replace and bypass relation between advices and process models. They did not make it clear if there it is possible to define parallel advices for a join point, i.e. M.R.A.4 requirement. If we asssume that they support parallel advice, there will be inconsistency in the approach when an advice wants to replace the join point, and another advice tries to bypass it at the same time. Thus, we have to assume that the authors did not consider this requirement, which result in lower degree of support for separation of cross-cutting concerns from process models in comparison to AO4BPMN (see Fig. 5 where this approach is a sub-set of the AO4BPMN).
4. Cappelli et al. [2] propose an extension to BPMN to support aspect oriented business process modeling. This approach has the same degree of separation as AOBPMN in Signature Exposure, Transformation Patterns, and Phases Support. It has the best degree of separation in Rule Composition dimension as it can be seen in Fig. 5, since it can support rule composition without considering dominant perspective, i.e. M.R.R.6. For evaluating the Advice Relations, it is not clear if this approach support around scenario or not. There is no clue about whether this scenario can be defined in this approach. Moreover, it is not clear if this approach support nested advices and precedence between advices. Therefore, this paper assumes that these requirements are not supported. Thus, this approach is not strong in defining Advice Relations.

5. Jalali et al. [9] propose an approach to support aspect oriented business process management in Yet Another Workflow Language (YAWL). The approach only supports task signature exposure like other approaches. However, it enables composition of pointcut rules based on both task and data perspectives and their combinations. At the same time, it considers task perspective as the dominant dimensions when composing rules. Therefore, it supports a better degree of separation in terms of rule composition in comparison with Charfi et al. [4], Jabeen et al. [6] and Patiniotakis et al. [13], but it is weaker than Cappelli et al. [2] in rule composition. This comparison can be clearly seen in Fig. 5 where Cappelli et al. [2] approach has a higher degree in Rule Composition. This approach also supports the same degree of advice relations like AOBPMN. However, it is a little better in Transformation Patterns since it supports data transformation between process instances and advices. Finally, it has the best degree for supporting separation of cross-cutting concerns in phases support, where it supports design, run, backward-forward adjustment and forward adjustment.

6 Conclusion and Future Works

This paper proposes an assessment framework for evaluating aspect oriented business process modeling approaches for the first time. The framework measures different approaches based on five dimensions, which are defined for evaluating the degree of separation of cross-cutting concerns from process models. Each dimension defines a set of requirements, which are required for aspect oriented business process modeling. The requirements are defined through studying literatures and by lessons, which have been learned through enacting aspect oriented business process models in case studies. The framework is applied for current approaches, and the result shows the strength and weakness of each approach. It also enables comparison between different approaches to understand which one supports better degree of separation of cross-cutting concerns from process models. It also makes the gaps in this area clear, which resulted in discovering directions for prospective works in aspect oriented business process modeling. The future works can be defined in five identified categories:

- **Signature Exposure:**
  – How the process signature can be defined in aspect oriented business process modeling. This investigation can enable definition of aspects like a security concern which must be performed before starting the purchase process.
  – How the data signature can be defined in aspect oriented business process modeling. This investigation can enable definition of aspects for data elements. For example, a security concern which must be performed when a customer account balance is going to be changed.
  – How the resource signature can be defined in aspect oriented business process modeling. This investigation can enable definition of aspects for resource elements. For example, a security concern when some special people in organization want to perform a task.

- **Rule Composition:**
  – How process data can be used when composing rules. The use of process data in rule composition can support better degree of separation. For example, many complex scenarios can be addressed such as having a security mechanism if the account balance which is changed by a task exceeds some limits in comparison with the process level data.
  – How resource data can be used in composing rules. This is a very important aspect, since it enables the definition of many other aspects like if resources are busy, request for additional resource. The aspect oriented business process modeling can support many complex scenarios by supporting resource perspective in composition of rules.
- **Advice Relations:**
  – How nested advice can be modeled is also an important issue to be investigated. This investigation enables capturing more real cases such as a logging concern for a security concern.
  – How precedence should be supported is an important issue that is also investigated by Jalali et al. [8]. This requirement is critical since cross-cutting concerns do not always have the same order.
- **Transformation Patterns:**
  – It is important to investigate how to transform process data between main processes and cross-cutting concerns and vice versa. This mechanism enables transformation of knowledge between the main process and its related advices.
  – It is also important to investigate how resource information should be transferred between different process instances, for example, between the main process and its advices. This investigation enables enforcement of different resource patterns in aspect oriented business process modeling like four-eye principle, segregation of duties, binding of duties, conflict of interest and need-to-know scenarios [1].
- **Phases Support:**
  – The approach proposed by Jalali et al. [9] already supports most of the phases support requirements except backward adjustment. Thus, it is important to investigate how cross-cutting concerns can be adjusted while their instances are running.

Furthermore, the combinations of these requirements can be considered as directions for future research, since the full degree of separation cannot be addressed while all the requirements are not fulfilled. The framework can also be extended to cover more requirements. Moreover, it can inspire to define evaluation frameworks for other aspect oriented approaches in other disciplines like programming.

Acknowledgement. I would like to appreciate professor Paul Johannesson for his valuable comments on this work.

References

1. Accorsi, R.: An approach to data-driven detective internal controls for process-aware information systems. Data Usage Management on the Web, 29 (2011)
2. Cappelli, C., Santoro, F., do Prado Leite, J., Batista, T., Medeiros, A., Romeiro, C.: Reflections on the modularity of business process models: The case for introducing the aspect-oriented paradigm. BPM Journal 16, 662–687 (2010)
3. Charfi, A., Mezini, M.: Ao4bpel: An aspect-oriented extension to bpel. World Wide Web 10(3), 309–344 (2007)
4. Charfi, A., Müller, H., Mezini, M.: Aspect-Oriented Business Process Modeling with AO4BPMN. In: Kühne, T., Selic, B., Gervais, M.-P., Terrier, F. (eds.) ECMFA 2010. LNCS, vol. 6138, pp. 48–61. Springer, Heidelberg (2010)
5. Collell, D.C.: Aspect-oriented modeling of business processes. Master's thesis, der Technischen Universitat Darmstadt, Darmstadt (2012)
6. Jabeen, A., Tariq, S., Farooq, Q., Malik, Z.: A lightweight aspect modelling approach for bpmn. In: 2011 IEEE 14th International Multitopic Conference (INMIC), pp. 255–260 (2011)
7. Jalali, A., Johannesson, P.: Multi-Perspective Business Process Monitoring. In: Nurcan, S., Proper, H.A., Soffer, P., Krogstie, J., Schmidt, R., Halpin, T., Bider, I. (eds.) BPMDS 2013 and EMMSAD 2013. LNBIP, vol. 147, pp. 199–213. Springer, Heidelberg (2013)
8. Jalali, A., Wohed, P., Ouyang, C.: Aspect oriented business process modelling with precedence. In: Mendling, J., Weidlich, M. (eds.) BPMN 2012. LNBIP, vol. 125, pp. 23–37. Springer, Heidelberg (2012)
9. Jalali, A., Wohed, P., Ouyang, C., Johannesson, P.: Dynamic weaving in aspect oriented business process management. In: Meersman, R., Panetto, H., Dillon, T., Eder, J., Bellahsene, Z., Ritter, N., De Leenheer, P., Dou, D. (eds.) ODBASE 2013. LNCS, vol. 8185, pp. 2–20. Springer, Heidelberg (2013)
10. Kiczales, G., Lamping, J., Mendhekar, A., Maeda, C., Lopes, C., Loingtier, J., Irwin, J.: Aspect-oriented programming. In: Akşit, M., Matsuoka, S. (eds.) ECOOP 1997. LNCS, vol. 1241, pp. 220–242. Springer, Heidelberg (1997)
11. La Rosa, M., Wohed, P., Mendling, J., ter Hofstede, A., Reijers, H., Van der Aalst, W.M.P.: Managing process model complexity via abstract syntax modifications. IEEE Transactions on Industrial Informatics 7(4), 614–629 (2011)
12. I. O. Object Management Group. Business Process Model and Notation (BPMN). Technical report, Object Management Group, Inc. (OMG) (2013)
13. Patiniotakis, I., Papageorgiou, N., Verginadis, Y., Apostolou, D., Mentzas, G.: An aspect oriented approach for implementing situational driven adaptation of bpmn2.0 workflows. In: La Rosa, M., Soffer, P. (eds.) BPM Workshops 2012. LNBIP, vol. 132, pp. 414–425. Springer, Heidelberg (2013)
14. Shankardass, A.: The dynamic adaptation of an aspect oriented business process in a service oriented architecture platform. Master's thesis, Athabasca University, Athabasca, Alberta (2009)
15. Wang, J., Zhu, J., Liang, H., Xu, K.: Concern oriented business process modeling. In: IEEE International Conference on e-Business Engineering, ICEBE 2007, pp. 355–358 (2007)

Aspect Mining in Business Process Management

Amin Jalali

Department of Computer and Systems Sciences,
Stockholm University
aj@dsv.su.se

Abstract. Automatic discovery of process models from event logs is an important and promising area in Business Process Management. Process models document how business processes should be performed, so they capture different concerns related to business processes. Some of these concerns are not limited to one process model, and they are repeated in many others as well, called cross-cutting concerns. Although many works have been done to enable discovering different process models, there is no investigation about how models with cross-cutting concerns can be discovered from even logs. Therefore, this work proposes an approach to enable discovering these models from event logs. The investigation is performed based on a case-study from the banking domain. The result shows how these concerns hinder existing process discovery techniques, and how the proposed approach can solve the problem.

Keywords: Process Mining, Aspect Mining, Process Discovery, Business Process Management, Aspect Oriented.

1 Introduction

Automatic process discovery is an important area of research in the Business Process Management (BPM), since it enables discovering process models from what really have happened rather than what people claim. The reality is recorded in event logs, so process discovery algorithms aim to find the process models from these logs, which results in discovering the models that have been really followed in a business, called de-facto models. Traditionally, processes are discovered through interviews and other techniques, which results in models that describe what should be done rather than what is really happening. These models are called de-jure models. The comparison between these models can help businesses to discover the deviations between the reality and the goals, which points out opportunities for re-engineering business processes. Therefore, it is very promising and important to discover process models from event logs in BPM area.

Process models document various aspects of a business process, which fulfil different concerns. For example, a process model might incorporate the definition of security concerns that should be followed in a business process. Charfi et al. [15] mention auditing, authorization, privacy, and many other concerns as complementary aspects, which should be considered when designing a process

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 246–260, 2014.

model. Therefore, a process model might contain aspects which are in common between different process models. These aspects are *scattered* in process models, and each process is *tangled* to them. The concerns which introduce the scattering and tangling relation to process models are called *cross-cutting concerns*. These concerns hinder maintenance and managing changes in process models. They also decrease the re-usability of realization of concerns in different process models. Moreover, they challenge enforcing process models to be comply with rules, since a change might be forgotten to be applied in a part of the process which reflects such a concern. Therefore, these concerns should be separated from process models [11,15,22].

To separate cross-cutting concerns from process models, Aspect Oriented Business Process Management (AOBPM) introduces a new modularization technique to encapsulate these concerns into modules. It defines how these modules should be related to each other to solve both scattering and tangling problem. As a result, models are less complex, more reusable and easier to maintain [11,15,22]. The models can be combined together in order to support the business process enactment, called weaving [23]. Currently, the whole BPM lifecycle is supported for AOBPM, which enables separation of cross-cutting concerns from business processes.

Although the complete BPM lifecycle is enabled by modeling and execution of aspect oriented business process models, the starting point is the design phase, where the designers model business processes using this approach. Indeed, it applies a lot of cost for organizations that already have process models, which are designed in a traditional way. Furthermore, it is not possible to completely discover such models from process logs using existing process mining techniques, since none of them can completely discover process models with duplicate activities [3] - which are scattered across one or different models. This limitation also hinders process mining area since it is common to have process models with activities which are repeated in a model, and this limits the area of application of process mining.

Therefore, this paper investigates how models with cross-cutting concerns can be discovered from even logs. It starts by investigating the problem through a banking case study in which it will be shown how current process mining techniques cannot discover process models that have cross-cutting concerns in section 2. Then, this paper proposes a method with which such process can be discovered in section 3. The method is later applied and demonstrated through the same case study in section 4. Moreover, a critical discussion on related works are given which show why those approaches cannot solve the problem in section 5. Finally, section 6 concludes the paper and shows the direction for future works.

2 Case Study

In this section, we can see how cross-cutting concerns in business processes hinders current process mining techniques when discovering process models from event logs. Thus, this section starts by introducing a dealing process of a bank.

Then, it shows how current techniques in process mining cannot discover this process based on its events log.

2.1 Case Description

This section describes the scenario of the selected case from the financial domain. The banking case was selected because there are many cross-cutting concerns in banking domain like securty, privacy, logging, etc. To choose appropriate processes, i.e. fairly simple yet representative processes with at least a couple of cross-cutting concerns, an interview with a domain expert from a bank was conducted . For the confidentiality reasons, the bank asked to be remained anonymous. Two processes were selected. Here, one of them is presented namely the *Change asset deal* process. Detailed information about the process was derived through a follow-up interview with the same domain expert.

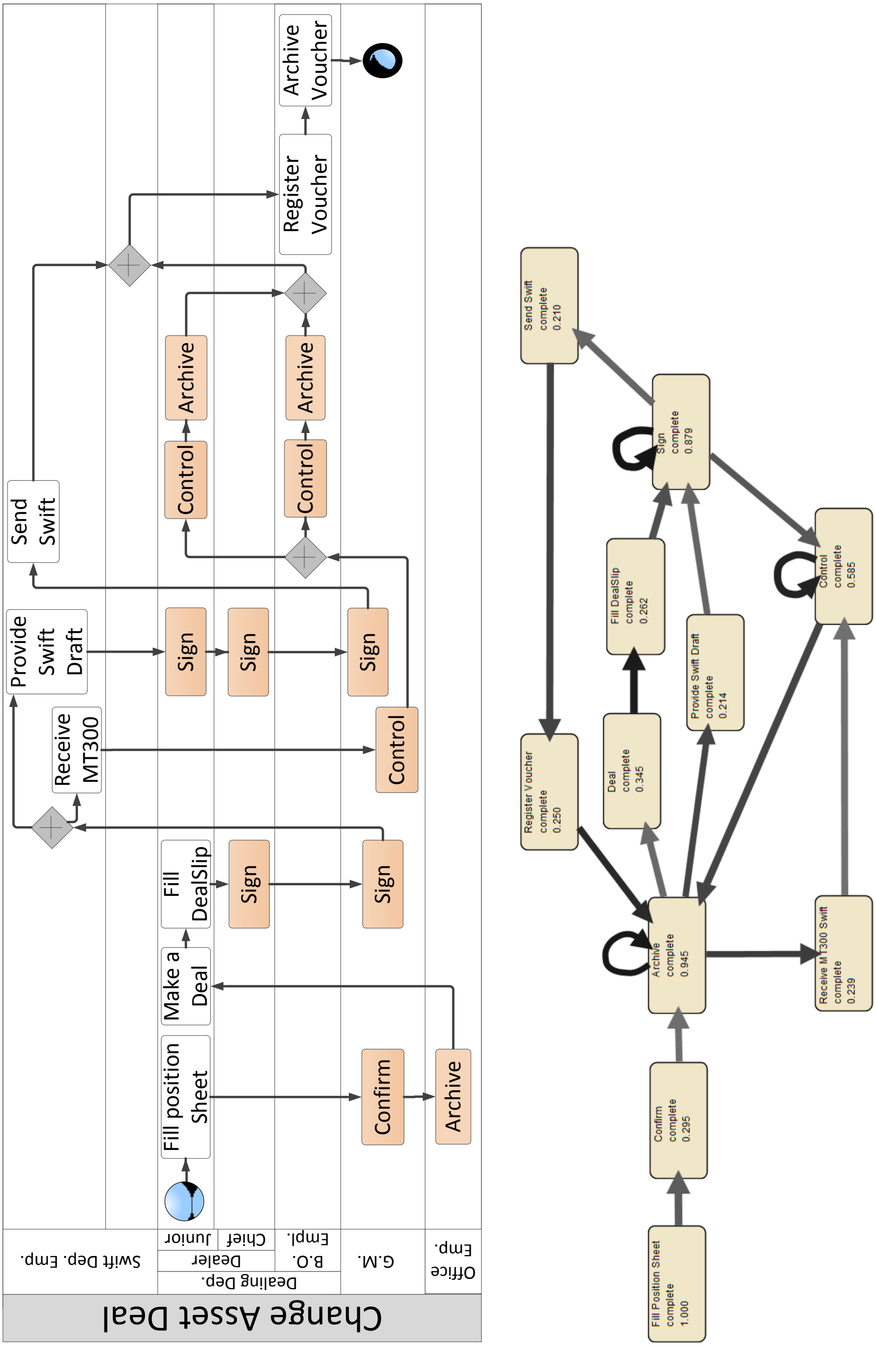
Generally, the assets of the bank are in two forms, cash and non-cash. Cash assets are either in the form of the account balances of the bank or the marketable securities. The `Change asset deal` process (see Figure 1(a)) handles deals for exchanging assets of the bank from one currency to another. The process starts with a *junior dealer* filling in a position sheet (`Fill position sheet` activity). Then, the *general manager* confirms the position sheet, which will be archived by the *office employee*. Then, a *junior dealer* makes the deal and fills in a deal slip. Next, both a *chief dealer* and the *general manager* sign the deal slip.

Afterwards, two parallel sets of activities are performed. On the one hand, the dealt amount of money is sent to the external partner of the deal. For this, first an *employee of the Swift department* provides a swift draft for sending the money. Then, for security purposes, the *junior dealer*, *chief dealer* and *general manager* sign the swift draft. Finally, an *employee of the Swift department* sends out the swift. On the other hand, the dealt amount of money should be received. This part starts when an *employee of the Swift department* receives an MT300 swift message. The *employee* sends this message to the general manager. The *general manager* makes an order to the Back office department and to the dealer to control the swift message. These messages are controlled separately. When each one of them has been controlled, the messages are archived (separately). When the deal is made, a *back office employee* registers a voucher in the accounting system. Finally, the deal is archived.

As it can be seen, the activities that represent cross-cutting concerns are repeated in this model. The events log which is the result of execution of this model has also repetitive activities. Next section shows the result of applying the current process mining algorithm on this log.

2.2 Discovering the Case Study Process

This section shows the result of applying two process mining algorithm on events log of the case study. The log was cleaned, so it is noise free. The log was also structured according to the eXtensible Event Stream (XES) standard, adopted by the IEEE Task Force on Process Mining [2]. In this standard, each event



(a) Change Asset Deal Process

(b) The process model discovered through applying Fuzzy Miner

Fig. 1. Case Study Result

represents the execution result of an activity. The sequences of events which are related to a specific case, i.e. the process instance, are grouped into a trace. The events in a trace are ordered according to their execution times. Therefore, the event log consists of number of traces that represent different execution scenarios of a process model.

Two process mining algorithms are selected to be applied on the log, i.e. Fuzzy algorithm and Alpha algorithm. Fuzzy algorithm is selected since it can discover process models from event logs with noise [18]. Alpha algorithm is used since it supports discovering process models from noise free events log [4].

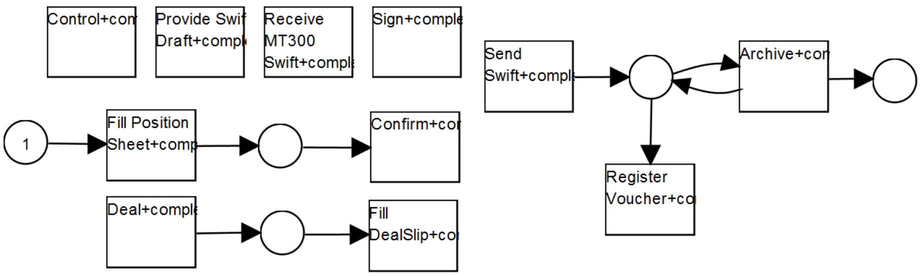


Fig. 2. The process model discovered through applying Alpha Miner

Fig. 1(b) shows the process model that is discovered through using Fuzzy Miner in ProM [17]. The fuzzy miner is the plug-in that implements the fuzzy algorithm to discover process models from event logs which might have noise. Although the result is a process model, it does not represent the correct process that generates the log (compare it with Fig. 1(a)). Indeed, the discovered process is not correct and the result is useless.

Fig. 2 shows the process model that is discovered through using Alpha Miner in ProM. The alpha miner is the plug-in that implements alpha algorithm to discover process models from a noise free event logs. As it can be seen, the discovered process is fragmented, and it cannot show the proper control-flow of the process.

In the next section, this paper proposes an approach with which process models with cross-cutting concerns can be discovered from event logs.

3 Approach

This section proposes an approach to enable discovering process models which contain cross-cutting concerns. The approach is a cyclical process which contains four phases, i.e. i) cross-cutting concerns discovery, ii) cross-cutting concerns elimination, iii) business process discovery, and iv) relation discovery (see Fig. 3).

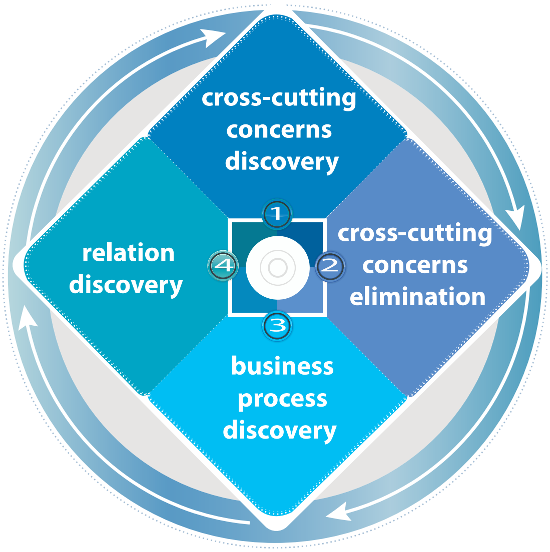


Fig. 3. Aspect Mining Approach for BPM

3.1 Cross-Cutting Concerns Discovery

In this phase, cross-cutting concerns should be discovered. These concerns can be discovered manually or automatically. In *manual discovery* of cross-cutting concerns, the concerns can be identified through traditional methods of process identification and discovery like the interview. For example, cross-cutting concerns can be expressed as policies in organizations. In *automatic discovery* of cross-cutting concerns, the concerns should be discovered by an algorithm. These concerns can be represented as events in the log file recording the enactment result of a process or several process models.

To discover the cross-cutting concerns from a log file of a process model, an algorithm should search for processes with duplicate tasks. Finding these tasks helps to identify cross-cutting concerns, since these concerns are sometimes repeated in a process model. Li et al [24], proposes an extension to alpha algorithm to discover repetitive activities in an event log, which is extended by Gu et al [27] to discover short loops. Their extensions do not work in all situations, yet it can help in some others. In our case study, this approach did not help to discover repetitive activities. Section 5 discusses the problem of this algorithm in more detail. Alves de Medeiros [26] proposes a genetic algorithm to tackle duplicate tasks, called Duplicates Genetic Algorithm (DGA). This algorithm helps to discover some processes with duplicate tasks automatically. Herbst et al. [21] also proposes an algorithm which can help in identifying duplicate tasks in process models. Moreover, there are many works in programming area which can be considered to inspire extending this area, e.g. [8,29].

To discover the cross-cutting concerns from a log file of multiple process models, no solution exist currently. This is due to current situation in process mining area in which "there are no process discovery techniques that produce overarching models able to relate and analyze different groups and process variants" [1]. However, it is expected that the introduction of process cubes, which support vertical and horizontal decomposition of event logs, enables such investigation [1].

3.2 Cross-Cutting Concerns Elimination

Identifying cross-cutting concerns enables separation of their events from event logs. Therefore, the logs will contain only the main activities for each process model. It should be mentioned that the whole approach is cyclical, so all concerns may be identified and their events may be eliminated through several cycles.

3.3 Business Process Discovery

The cleaned events log can be used by existing process mining algorithms to discover process models. There are many algorithms that can be used, for example alpha algorithm can be used to identify process models from noise free events logs [4], and fuzzy miner can be used to discover process models from events logs which have noises [18].

3.4 Relation Discovery

The relation between cross-cutting concerns and process models can be discovered manually or automatically. These relations are called *pointcuts* in aspect oriented paradigm.

In *manual discovery*, pointcuts can be identified through traditional methods like the interview. In *manual discovery*, pointcuts should be identified and added to process models through *process enhancement* technique. Process Enhancement is a technique "to extend or improve an existing process model using information about the actual process recorded in some event log" [2]. Here, information about the control-flow perspective of the model and the cross-cutting concerns are available as a result of previous phases. Moreover, complete information regarding the execution of these events exists in the event log. Therefore, the relation between the process model and cross-cutting concerns can be identified through investigating the event logs. This information can be added to the process model as pointcuts which defines these relations. Fig. 4 shows the pointcut discovery phases.

4 Demonstration

The proposed approach is applied on event logs of the case study. The logs are noise free and cleaned. The evaluation is performed in four phases, i.e. i) cross-cutting concerns discovery, ii) cross-cutting concerns elimination, iii) business process discovery, and iv) relation discovery.

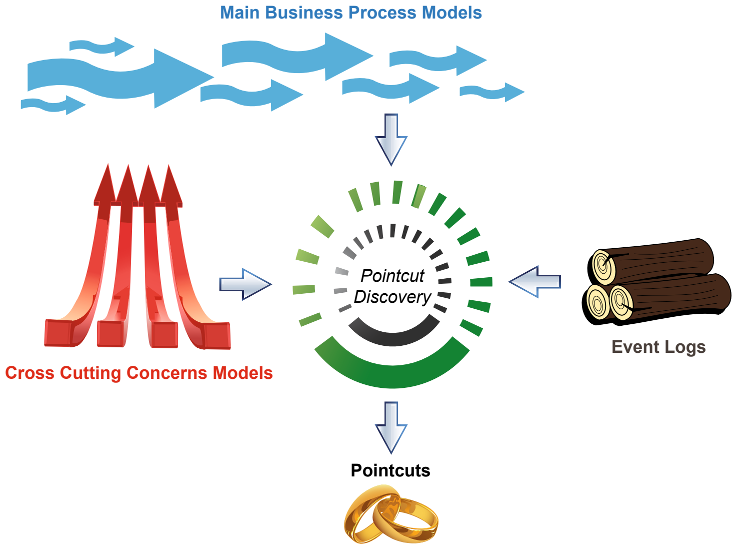


Fig. 4. Pointcut Discovery Process

4.1 Cross-Cutting Concerns Discovery

To discover cross-cutting concerns, manual cross-cutting concerns discovery is followed through interview, because there is not currently any algorithm that can be used for automatic discovery of aspects. There are four scenarios in this business process that require some sort of security checks, i.e. i) the position sheet should be confirmed before making a deal, ii) the deal slip should be signed for further processing, iii) the received swift message should be controlled, and iv) the swift draft should be signed before it can be sent.

These four scenarios can be defined through three mechanisms that expressed three cross-cutting concerns, i.e. `Sign`, `Confirm` and `Control`. As the `Sign` mechanism, if the document is a swift draft, the junior dealer, chief dealer and general manager should sign it sequentially; however, if it is a dealslip, only chief dealer and general manager require to sign it. This concern is modeled as a process model which is called Sign (see Fig. 5). As the `Confirm` mechanism, the general manager can only confirm decisions, and each confirmation should be archived by the office employee. This concern is modeled as a process model which is called Confirm (see Fig. 5). As the `Control` mechanism, the general manager should control all swifts which are received. Then, the swifts message should be controlled and archived by the back-office employee. If the received swift is related to a deal, the dealer who imitated the deal should also control and archive the swift. This concern is modeled as a process model which is called Control (see Fig. 5). All of these concerns represent security aspects of the process model, so they are grouped together in a module, called *security aspect*.

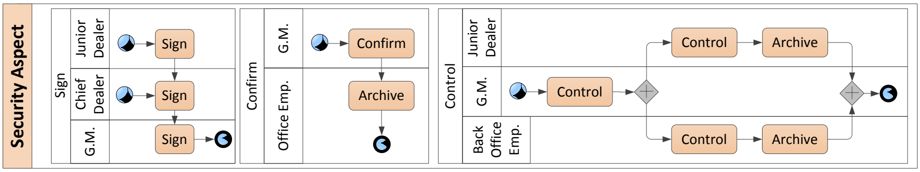


Fig. 5. The result of manual discovery of cross-cutting concerns

4.2 Cross-Cutting Concerns Elimination

In this step, all events which are related to one of the activities in cross-cutting concerns are removed from event log. Therefore, the new events log only contains information about the core part of the change asset deal process.

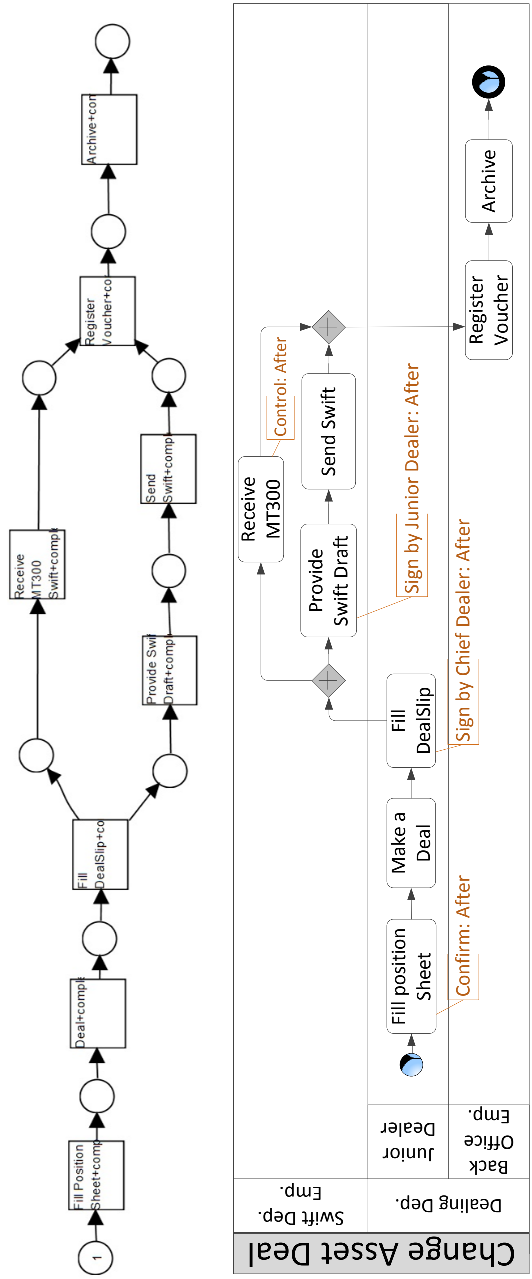
4.3 Business Process Discovery

The cross-cutting concerns free event log is used to discover the change asset deal process. Fig. 6(a) shows the result of process discovery by applying Alpha Miner in ProM. The same result was achieved by applying Fuzzy Miner. As it was expected, the results are the same since the log file does not contain noises. The discovered process represents the change asset deal process without cross-cutting concerns. It starts when a junior dealer fills a position sheet. Then, (s)he should make a deal, and fill the deal slip. Afterwards, the swift department employee provides the swift draft and sends the swift. In parallel, the swift department of the corresponding bank sends the related swift, so the swift department employee also receives a swift with MT300 message. Finally, the back-office employee should register the voucher and archive it.

4.4 Relation Discovery

We already discovered both the deal asset process model and models of cross-cutting concerns. However, it is not clear how these models should be related to each other. Therefore, manual pointcut discovery is used to discover the relations. It should be noted that automatic pointcut discovery is not possible now, since it requires further research to define an algorithm for such a discovery. Some of these relations are specified when we discovered cross-cutting concerns, e.g. the general manager should control all swifts which are received. Thus, in this step all relations should be investigated and completed.

Fig. 6(b) shows the discovered change asset deal process, in which the relations to cross-cutting concerns are specified through annotations. As it can be seen, the position sheet should be confirmed after junior dealer fills it. Moreover, the dealslip should be signed after junior dealer fills it. The swift draft should be signed by junior dealer, chief dealer and general manager before the swift is sent, and the received swift should be controlled. This process and cross-cutting concerns which are shown in Fig. 5 complement each other.



(a) The process model discovered through applying Alpha Miner

(b) Change Asset Deal process with pointcuts

Fig. 6. Demonstrating the proposed approach in the case study

5 Related Works and Discussions

This section explains related works in aspect mining through three categories, i.e. discovering duplicate tasks from events log in BPM, discovering cross-cutting concerns from business process models, and aspect mining in programming. The first category explains current attempts to discover duplicate tasks from events logs in BPM area, and it explains a counter example for these works. The second category introduces some works which enable discovering cross-cutting concerns from business process models. The third category shows the similar paradigm in programming area which tries to discover cross-cutting concerns from programs' codes and traces.

5.1 Discovering Duplicate Tasks from Event Logs in BPM

Discovering duplicate tasks is one of the challenges and open issues in process mining area [3,25]. Li et al [24], propose an extension to alpha algorithm to discover duplicate tasks in events logs, which is also extended by Gu et al [27] to discover short loops. This algorithm is called $\alpha*$. However, their extensions do not work in all situations. It cannot discover duplicate activities in our case study correctly, which can be consider as a counter example.

Their approach follows two steps, i.e. i) finding duplicate tasks, and ii) discovering a process model. To find duplicate tasks, the approach makes a predecessor and successor table, called P/S-table [24]. Each trace is denoted by δ_i, where i represents the number of trace that contains events. $\delta_i(t, n)$ represents a function which specify the nth occurrence of event t in δ_i. For each of these occurrence, the predecessor and successor of the event can be retrieved by T_P and T_S correspondingly. The approach proposes three rules with which the duplicate tasks claimed to be discovered from event logs, i.e. Rule (1), (2) and (3). The U in these formula represents the set of duplicate tasks.

$$\text{(1) IF } ((T_P \neq T_{P'}) \text{ AND } (T_S \neq T_{S'}) \text{ AND } (T_P \neq T_{S'}) \text{ AND } (T_S \neq T_{P'}) \\ \text{AND (not } T_P \Sigma w T_{P'})) \text{ THEN } <\delta_i(t, N_1), \delta_j(t, N_2)> \in \text{U}$$

$$\text{(2) IF } ((T_P = T_{S'}) \text{ AND } ((T_{P'} \neq \mathit{pred}(\delta_i, T_P)) \text{ OR } (T_S \neq \mathit{succ}(\delta_j, T_{S'}))) \\ \text{THEN } <\delta_i(t, N_1), \delta_j(t, N_2)> \in \text{U}$$

$$\text{(3) IF } ((T_S = T_{P'}) \text{ AND } ((T_{P'} \neq \mathit{pred}(\delta_j, T_{P'})) \text{ OR } (T_{S'} \neq \mathit{succ}(\delta_i, T_S))) \\ \text{THEN } <\delta_i(t, N_1), \delta_j(t, N_2)> \in \text{U}$$

The case study in this paper shows that this formula cannot work. To explain it, we can consider a very simple process shown in Fig. 7 which contains tasks a,b,c,a and b which should be followed in sequence. It is clear that all traces of executing this process has the same occurrence of events, so for each i and j representing a trace number: $\delta_i = \delta_j$.

The duplicate task a cannot be discovered through the first rule since:

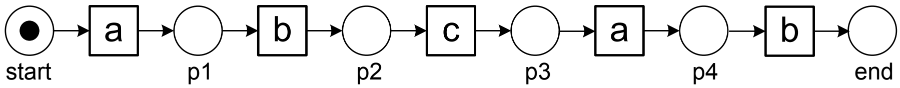


Fig. 7. Counter Example for duplicate task identification rules in α*-algorithm

$$\frac{\begin{array}{c}\text{For rule 1:}\\ T_S = \delta_i(a,1) = b\\ T_{S'} = \delta_j(a,2) = b\end{array}}{\therefore T_S = T_{S'} \nRightarrow a \in U}$$

The duplicate task a cannot be discovered through the second rule since:

$$\frac{\begin{array}{c}\text{For rule 2:}\\ T_P = \delta_i(a,1) = \emptyset\\ T_{S'} = \delta_j(a,2) = b\end{array}}{\therefore T_P \neq T_{S'} \nRightarrow a \in U}$$

The duplicate task a cannot be discovered through the third rule since:

$$\frac{\begin{array}{c}\text{For rule 3:}\\ T_S = \delta_i(a,1) = b\\ T_{P'} = \delta_j(a,2) = c\end{array}}{\therefore T_S \neq T_{P'} \nRightarrow a \in U}$$

Thus, this algorithm cannot discover duplicate tasks in all circumstances.

5.2 Discovering Cross-Cutting Concerns from Process Models

Although cross-cutting concerns can be discovered from events log, they can also discovered from process models that are designed through traditional modelling techniques. There are few works in BPM area that support separation of cross-cutting concerns from process models,e.g. [16,5,6]. These works apply queries on process models definitions, and they try to discover cross-cutting concerns which exist within these models.

5.3 Aspect Mining in Programming

There are a lot of researches about Aspect Mining in Aspect Oriented Programming area, e.g. [7,10,13,14,20]. These researches can be divided into two categories, i.e. static and dynamic program analysis techniques.

Static program analysis techniques aim to discover aspects out of source codes. Two approaches exist in these techniques, i.e. discovering advices without any previous knowledge or with previous knowledge. For example, 'Formal Concept Analysis of Identifiers' is utilized by some researchers to discover cross-cutting concerns from Smalltalk and Java codes using the FCA algorithm [12,31]. These techniques require that the user have some assumption regarding the structure of advices. Other researchers used 'Natural Language Processing on Source Code' to solve this problem [28]. Furthermore, there are other works that discover cross-cutting concerns from codes without any previous knowledge about concerns [19]. There are also other techniques based on detecting clones in codes, which are categorized in these techniques [19]. The advantages of static approaches are that they consider all possible execution scenarios in source codes.

Dynamic program analysis techniques aim to discover aspects out of program traces of method invocations. There are also different works that have been done in this area. For example, discovering the cross-cutting concerns by detecting recurring execution patterns [9], or applying formal concept analysis to achieve this goal [30]. These works can be used as inspiration for enabling automatic cross-cutting concerns discovery in BPM area.

6 Conclusion

This paper proposes an approach that enables discovering business process models which have cross-cutting concerns. The approach is a cyclical process which has four phases, i.e. i) cross-cutting concerns discovery, ii) cross-cutting concerns elimination, iii) business process discovery, and iv) relation discovery. In *cross-cutting concerns discovery*, the cross-cutting concerns should be discovered either manually or automatically. The events of these concerns should be eliminated from event log in the *cross-cutting concerns elimination* phase. The eliminated log can be used for discovering the main process model in *business process discovery* phase. Finally, the relations between cross-cutting concerns and the main process model should be discovered in *relation discovery* phase, which can be done manually or automatically. This approach is a cyclical process, which means that it might be repeated several times that all cross-cutting concerns are separated.

The approach is applied and demonstrated through a case study, which shows how these sort of models can be discovered. The finding is important since it is common to have repetitive and duplicate tasks in business processes, and this situation is a limitation of process mining currently. It means that process mining cannot deal with these sorts of processes, yet this approach can enable discovering these processes as well.

As future works, more experiment are needed to be conducted to evaluate the proposed method. Moreover, it is important to investigate i) how cross-cutting concerns can be discovered automatically from event logs, ii) how duplicate tasks can be discovered automatically from event logs, iii) how the relations between cross-cutting concerns and the main process model can be discovered automatically from the event log.

Acknowledgement. I would like to appreciate professor Paul Johannesson for his valuable comments on this work.

References

1. van der Aalst, W.M.P.: Process cubes: Slicing, dicing, rolling up and drilling down event data for process mining. In: Song, M., Wynn, M.T., Liu, J. (eds.) AP-BPM 2013. LNBIP, vol. 159, pp. 1–22. Springer, Heidelberg (2013)
2. van der Aalst, W., et al.: Process mining manifesto. In: Daniel, F., Barkaoui, K., Dustdar, S. (eds.) BPM Workshops 2011, Part I. LNBIP, vol. 99, pp. 169–194. Springer, Heidelberg (2012)
3. van der Aalst, W.M.P., Weijters, A.J.M.M.: Process mining: a research agenda. Computers in Industry 53(3), 231–244 (2004)
4. van der Aalst, W.M.P., Weijters, T., Maruster, L.: Workflow mining: discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering 16(9), 1128–1142 (2004)
5. Beheshti, S.M.R.: Organizing, Querying, and Analyzing Ad-hoc Processes Data. PhD thesis, University of New South Wales Sydney, Australia (2012)
6. Beheshti, S.-M.-R., Benatallah, B., Motahari-Nezhad, H.R.: Enabling the analysis of cross-cutting aspects in ad-hoc processes. In: Salinesi, C., Norrie, M.C., Pastor, Ó. (eds.) CAiSE 2013. LNCS, vol. 7908, pp. 51–67. Springer, Heidelberg (2013)
7. Breu, S.: Extending dynamic aspect mining with static information. In: Fifth IEEE International Workshop on Source Code Analysis and Manipulation 2005, pp. 57–65 (2005)
8. Breu, S., Krinke, J.: Aspect mining using event traces. In: Proceedings of the 19th International Conference on Automated Software Engineering 2004, pp. 310–315 (September 2004)
9. Breu, S., Krinke, J.: Aspect mining using event traces. In: Proceedings of the 19th International Conference on Automated Software Engineering 2004, pp. 310–315. IEEE (2004)
10. Breu, S.: Th. Zimmermann. Mining aspects from version history. In: 21st IEEE/ACM International Conference on Automated Software Engineering, ASE 2006, pp. 221–230. IEEE (2006)
11. Cappelli, C., Santoro, F.M., do Prado Leite, J.C.S., Batista, T., Medeiros, A.L., Romeiro, C.S.C.: Reflections on the modularity of business process models: The case for introducing the aspect-oriented paradigm. BPM Journal 16, 662–687 (2010)
12. Ceccato, M., Marin, M., Mens, K., Moonen, L., Tonella, P., Tourwe, T.: A qualitative comparison of three aspect mining techniques. In: Proceedings of the 13th International Workshop on Program Comprehension, IWPC 2005, pp. 13–22. IEEE (2005)
13. Ceccato, M., Marin, M., Mens, K., Moonen, L., Tonella, P., Tourwé, T.: Applying and combining three different aspect mining techniques. Software Quality Journal 14(3), 209–231 (2006)
14. Ceccato, M., Marin, M., Mens, K., Moonen, L., Tonella, P., Tourwé, T.: Applying and combining three different aspect mining techniques. Software Quality Journal 14(3), 209–231 (2006)
15. Charfi, A., Müller, H., Mezini, M.: Aspect-Oriented Business Process Modeling with AO4BPMN. In: Kühne, T., Selic, B., Gervais, M.-P., Terrier, F. (eds.) ECMFA 2010. LNCS, vol. 6138, pp. 48–61. Springer, Heidelberg (2010)
16. Di Francescomarino, C., Tonella, P.: Crosscutting concern mining in business processes. IET Software 5(6), 552–562 (2011)

17. van Dongen, B.F., de Medeiros, A.K.A., Verbeek, H.M.W(E.), Weijters, A.J.M.M.T., van der Aalst, W.M.P.: The prom framework: A new era in process mining tool support. In: Ciardo, G., Darondeau, P. (eds.) ICATPN 2005. LNCS, vol. 3536, pp. 444–454. Springer, Heidelberg (2005)
18. Günther, C.W., van der Aalst, W.M.P.: Fuzzy mining – adaptive process simplification based on multi-perspective metrics. In: Alonso, G., Dadam, P., Rosemann, M. (eds.) BPM 2007. LNCS, vol. 4714, pp. 328–343. Springer, Heidelberg (2007)
19. Gybels, K., Kellens, A.: An experiment in using inductive logic programming to uncover pointcuts. In: First European Interactive Workshop on Aspects in Software (2004)
20. He, L., Bai, H.: Aspect mining using clustering and association rule method. International Journal of Computer Science and Network Security 6(2A), 247–251 (2006)
21. Herbst, J., Karagiannis, D.: Workflow mining with inwolve. Computers in Industry 53(3), 245–264 (2004); Process / Workflow Mining
22. Jalali, A., Wohed, P., Ouyang, C.: Aspect oriented business process modelling with precedence. In: Mendling, J., Weidlich, M. (eds.) BPMN 2012. LNBIP, vol. 125, pp. 23–37. Springer, Heidelberg (2012)
23. Jalali, A., Wohed, P., Ouyang, C., Johannesson, P.: Dynamic weaving in aspect oriented business process management. In: Meersman, R., Panetto, H., Dillon, T., Eder, J., Bellahsene, Z., Ritter, N., De Leenheer, P., Dou, D. (eds.) ODBASE 2013. LNCS, vol. 8185, pp. 2–20. Springer, Heidelberg (2013)
24. Li, J., Liu, D., Yang, B.: Process mining: Extending α-algorithm to mine duplicate tasks in process logs. In: Chang, K.C.-C., Wang, W., Chen, L., Ellis, C.A., Hsu, C.-H., Tsoi, A.C., Wang, H. (eds.) APWeb/WAIM 2007. LNCS, vol. 4537, pp. 396–407. Springer, Heidelberg (2007)
25. de Medeiros, A.K.A., van der Aalst, W.M.P., Weijters, A.J.M.M.T.: Workflow mining: Current status and future directions. In: Meersman, R., Schmidt, D.C. (eds.) CoopIS/DOA/ODBASE 2003. LNCS, vol. 2888, pp. 389–406. Springer, Heidelberg (2003)
26. de Medeiros, A.K.A.: Genetic process mining. PhD thesis, Technische Universiteit Eindhoven (2006)
27. Gu, C.Q., Chang, H.Y., Yi, Y.: Workflow mining: Extending the α-algorithm to mine duplicate tasks. In: 2008 International Conference on Machine Learning and Cybernetics, vol. 1, pp. 361–368 (2008)
28. Shepherd, D., Pollock, L., Tourwé, T.: Using language clues to discover crosscutting concerns. ACM SIGSOFT Software Engineering Notes 30, 1–6 (2005)
29. Tonella, P., Ceccato, M.: Aspect mining through the formal concept analysis of execution traces. In: Proceedings of the 11th Working Conference on Reverse Engineering 2004, pp. 112–121 (November 2004)
30. Tonella, P., Ceccato, M.: Aspect mining through the formal concept analysis of execution traces. In: Proceedings of the 11th Working Conference on Reverse Engineering 2004, pp. 112–121. IEEE (2004)
31. Tourwe, T., Mens, K.: Mining aspectual views using formal concept analysis. In: Fourth IEEE International Workshop on Source Code Analysis and Manipulation 2004, pp. 97–106. IEEE (2004)

Discovering Metric Temporal Business Constraints from Event Logs

Fabrizio Maria Maggi

University of Tartu, Estonia
f.m.maggi@ut.ee

Abstract. Process discovery aims at building process models using information retrieved from logs. Process characteristics play a significant role in the selection of a suitable process modeling language for describing process discovery results. Business processes characterized by high variability, in which participants have a lot of autonomy and flexibility in executing the process, are difficult to be described with procedural process modeling languages, since they explicitly represent in a model every possible path. Declarative languages, like Declare, alleviate this issue by defining a set of constraints between activities that must not be violated during the process execution instead of describing what to do step by step. Recently, several process discovery techniques have been proposed for extracting a set of Declare constraints from a log. However, no one of these techniques allows the user to exploit the time perspective often available in a log to discover "time-aware" Declare constraints. *Timed Declare* has already previously been introduced to monitor metric temporal constraints at runtime. In this paper, we use this semantics for discovering a set of Timed Declare constraints from an event log. We have implemented the proposed approach as a plug-in of the process mining tool ProM. We have validated the approach by using our plug-in to mine two real-life event logs.

Keywords: Process Discovery, Metric Temporal Logic, Event Correlations, Timed Declare.

1 Introduction

Process discovery is one of the three branches of the family of process mining techniques together with *conformance checking* and *process enhancement*. Through process discovery, it is possible to build from scratch a process model describing the behavior of a business process as recorded in an event log. Recently, several works have investigated advantages and disadvantages of using procedural or declarative process modeling languages to describe the results of a process discovery technique [13,14]. The results of these studies highlighted that the dichotomy *procedural versus declarative* reflects the nature of the process under examination. Procedural models like Petri nets, BPMN, and EPCs are more suitable to support business processes working in stable environments, in which participants have to follow predefined procedures, since they suggest step

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 261–275, 2014.

by step what to do next. In contrast, declarative process modeling languages, like Declare, provide process participants with a (preferably small) set of rules to be followed during the process execution. In this way, process participants have the flexibility to follow any path that does not violate these rules.

Declare is a declarative language introduced in [2] that combines a formal semantics grounded in Linear Temporal Logic (LTL) with a graphical representation for users.[1] A *Declare map* is a set of Declare constraints each one with its own graphical representation and LTL semantics (see [2] for a full overview of Declare). Recently, several process discovery techniques have been proposed for describing with a set of Declare constraints the behavior of a process as recorded in an event log [6,7,11,10]. However, no one of these techniques allows the user to extract "time-aware" Declare constraints from a log.

During the execution of a business process it is often extremely important to meet deadlines and optimize response times. To this aim, a Declare map can also include metric temporal constraints to guarantee the correct execution of a process in terms of *latencies* (related to events that cannot occur *before* a certain time, or must occur *after* a certain time) and *deadlines* (related to events that cannot occur *after* a certain time, or must occur *before* a certain time).

Timed Declare has already previously been introduced to monitor metric temporal constraints at runtime [16]. In this paper, we introduce an approach for discovering a set of Timed Declare constraints that are satisfied in a given log. When evaluating the satisfaction of a Timed Declare constraint, one often faces ambiguities in connecting events that "activate" the constraint (activations) and events that "fulfill" it (target events), since the activation of a constraint may potentially be associated to multiple target events. For example, consider traces $\mathbf{T}_1 = \langle a, b, c, b \rangle$ and $\mathbf{T}_2 = \langle a, a, b, b \rangle$ and the *response* constraint $\mathbf{G}(a \rightarrow \mathbf{F}b)$, meaning that if *a occurs*, then eventually *b follows* after a. It is unclear whether to associate the activation a of the constraint at $\mathbf{T}_1(1)$ with the occurrence of the target b at $\mathbf{T}_1(2)$ or $\mathbf{T}_1(4)$. We face similar ambiguity for the two activations of a in $\mathbf{T}_2$. If we want to discover Timed Declare constraints, we need to evaluate the time difference between the occurrence of a and its corresponding target b. A conservative approach, where we associate an occurrence of a with the closest occurrence of b could negatively affect the discovery results and lead to incorrect conclusions. Our proposed technique allows the user to guide the discovery task through event correlations to correctly evaluate the metric temporal constraints and to improve the quality of the discovered models.

We have implemented our proposed approach in a plug-in of the process mining tool ProM.[2] We have validated our discovery technique by using two real-life logs provided for the 2011 and 2012 BPI challenges [1,8] pertaining to the treatment of patients diagnosed with cancer in a large Dutch academic hospital and to a financial process in a Dutch financial institute.

The remainder of this paper is organized as follows. Section 2 presents some preliminaries about Declare, Timed Declare and event correlations. Section 3

[1] In the remainder, LTL refers to the version of LTL tailored towards finite traces.
[2] `www.processmining.org`

presents our algorithms for discovering Timed Declare constraints. Section 4 presents the validation of the approach. Finally, Section 5 concludes the paper.

2 Preliminaries

In this section, we introduce some preliminary notions. In particular, in Section 2.1, we give an overview of the Declare language. In Section 2.2 we introduce Timed Declare. In Section 2.3, we give a categorization of the correlation mechanisms used in this paper.

2.1 Declare

Declare is a declarative process modeling language introduced by Pesic and van der Aalst in [2]. A *Declare map* consists of a set of constraints which, in turn, are based on templates. Templates are parameterized classes of rules and constraints are their concrete instantiations. Here, we indicate template parameters with capital letters (see Tables 1 and 2) and real activities in their instantiations with lower case letters (e.g., constraint $\mathbf{G}(a \rightarrow \mathbf{F}b)$). Templates have a user-friendly graphical representation understandable to the user and their semantics are specified through LTL formulas. Each constraint inherits the graphical representation and semantics from its template.

For the sake of readability, in Tables 1 and 2, we use PLTL to define the semantics of Declare constraints, i.e., LTL augmented with past operators. The LTL rules used in this paper are constructed from propositional atoms by applying the future temporal operators $\mathbf{X}$ (next), $\mathbf{F}$ (future), $\mathbf{G}$ (globally), and $\mathbf{U}$ (until) in addition to the usual boolean connectives. Given a formula φ, $\mathbf{X}\varphi$ means that the next time instant exists and φ is true in the next time instant (strong next). $\mathbf{F}\varphi$ indicates that φ is true sometimes in the future. $\mathbf{G}\varphi$ means that φ is true always in the future. $\varphi\mathbf{U}\psi$ indicates that φ has to hold at least until ψ holds and ψ must hold in the current or in a future time instant.

PLTL extends LTL by introducing past operators. The past operators we use in this work are $\mathbf{Y}$ (yesterday), $\mathbf{O}$ (once), and $\mathbf{S}$ (since), which correspond to the future operators $\mathbf{X}$, $\mathbf{F}$, and $\mathbf{U}$ respectively. At any non-initial time, $\mathbf{Y}\varphi$ is true if and only if φ holds at the previous time instant. $\mathbf{O}\varphi$ indicates that φ is true at some past time instant (including the present time). $\psi\mathbf{S}\varphi$ is true if ψ holds somewhere in the past and φ is true from then up to now.

Tables 1 and 2 show the PLTL semantics of the Declare constraints used in this paper. Consider, for example, the *response* constraint $\mathbf{G}(a \rightarrow \mathbf{F}b)$. This constraint indicates that if a occurs, b must eventually follow. Therefore, this constraint is satisfied for traces such as $\mathbf{T}_1 = \langle a, b, c, b\rangle$, $\mathbf{T}_2 = \langle a, a, b, b\rangle$, and $\mathbf{T}_3 = \langle b, b, c, d\rangle$, but not for $\mathbf{T}_4 = \langle a, b, a, c\rangle$ because, in this case, the second a is not followed by a b. Note that, in $\mathbf{T}_3$, the *response* constraint is satisfied in a trivial way because a never occurs. In this case, we say that the constraint is *vacuously satisfied* [9]. In [5], the authors introduce the notion of *behavioral vacuity detection* according to which a constraint is non-vacuously satisfied in a

Table 1. Semantics and graphical notation for positive relation constraints

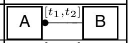
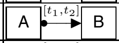
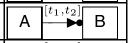
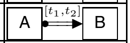
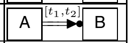
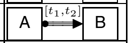
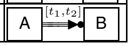
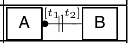
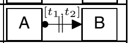
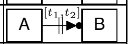
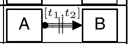
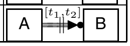
constraint	untimed semantics	timed semantics	notation
responded existence	$\mathbf{F}A \rightarrow \mathbf{F}B$	$\mathbf{G}(A \rightarrow (\mathbf{O}_{[t_1,t_2]}B \vee \mathbf{F}_{[t_1,t_2]}B))$	
response	$\mathbf{G}(A \rightarrow \mathbf{F}B)$	$\mathbf{G}(A \rightarrow \mathbf{F}_{[t_1,t_2]}B)$	
precedence	$\mathbf{G}(B \rightarrow \mathbf{O}A)$	$\mathbf{G}(B \rightarrow \mathbf{O}_{[t_1,t_2]}A)$	
alternate response	$\mathbf{G}(A \rightarrow \mathbf{X}(\neg A \mathbf{U} B))$	$\mathbf{G}(A \rightarrow \mathbf{X}(\neg A \mathbf{U}_{[t_1,t_2]} B))$	
alternate precedence	$\mathbf{G}(B \rightarrow \mathbf{Y}(\neg B \mathbf{S} A))$	$\mathbf{G}(B \rightarrow \mathbf{Y}(\neg B \mathbf{S}_{[t_1,t_2]} A))$	
chain response	$\mathbf{G}(A \rightarrow \mathbf{X}B)$	$\mathbf{G}(A \rightarrow \mathbf{X}_{[t_1,t_2]}B)$	
chain precedence	$\mathbf{G}(B \rightarrow \mathbf{Y}A)$	$\mathbf{G}(B \rightarrow \mathbf{Y}_{[t_1,t_2]}A)$	

Table 2. Semantics and graphical notation for negative relation constraints

constraint	untimed semantics	timed semantics	notation
not responded existence	$\mathbf{F}A \rightarrow \neg\mathbf{F}B$	$\mathbf{G}(A \rightarrow (\neg\mathbf{O}_{[t_1,t_2]}B \wedge \mathbf{F}_{[t_1,t_2]}B))$	
not response	$\mathbf{G}(A \rightarrow \neg(\mathbf{F}B))$	$\mathbf{G}(A \rightarrow \neg(\mathbf{F}_{[t_1,t_2]}B))$	
not precedence	$\mathbf{G}(B \rightarrow \neg(\mathbf{O}A))$	$\mathbf{G}(B \rightarrow \neg(\mathbf{O}_{[t_1,t_2]}A))$	
not chain response	$\mathbf{G}(A \rightarrow \neg(\mathbf{X}B))$	$\mathbf{G}(A \rightarrow \neg(\mathbf{X}_{[t_1,t_2]}B))$	
not chain precedence	$\mathbf{G}(B \rightarrow \neg(\mathbf{Y}A))$	$\mathbf{G}(B \rightarrow \neg(\mathbf{Y}_{[t_1,t_2]}A))$	

trace if the trace contains at least one activation of the constraint. An *activation* of a constraint in a trace is an event whose occurrence imposes, because of that constraint, some obligations on other events in the same trace. For example, a is an activation for the *response* constraint $\mathbf{G}(a \rightarrow \mathbf{F}b)$, because the execution of a forces b to be executed eventually.

In [7,11,10], algorithms for the discovery of Declare maps from event logs have been presented. In these works different notions of constraint support have been proposed. In this paper, we assume that the support of a constraint in a log is the percentage of traces in the given log in which the constraint is *non-vacuously satisfied*.

2.2 Timed Declare

We use Metric Temporal Logic (MTL) to define the semantics of Timed Declare constraints. We deal with a fragment of MTL where all traces are finite [15]. In Tables 1 and 2, we use the MTL future operators $\mathbf{X}_{[t_1,t_2]}$, $\mathbf{F}_{[t_1,t_2]}$, and $\mathbf{U}_{[t_1,t_2]}$. In addition we use the MTL past operators $\mathbf{Y}_{[t_1,t_2]}$, $\mathbf{O}_{[t_1,t_2]}$, and $\mathbf{S}_{[t_1,t_2]}$. Given a formula φ and the current time instant t, $\mathbf{X}_{[t_1,t_2]}\varphi$ means that the next time instant exists and falls into the interval $[t+t_1, t+t_2]$, and φ is true in the next time instant. $\mathbf{F}_{[t_1,t_2]}\varphi$ indicates that φ is true sometimes in the future in a time instant belonging to the interval $[t+t_1, t+t_2]$. $\varphi\mathbf{U}_{[t_1,t_2]}\psi$ indicates that φ has to hold at least until ψ holds and ψ must hold in a time instant belonging to the interval $[t+t_1, t+t_2]$. $\mathbf{Y}_{[t_1,t_2]}\varphi$ is true if φ holds at the previous time instant and this instant belongs to $[t-t_2, t-t_1]$. $\mathbf{O}_{[t_1,t_2]}\varphi$ indicates that φ is true at some past time instant (including the present time) falling into the interval $[t - t_2, t - t_1]$. $\psi\mathbf{S}_{[t_1,t_2]}\varphi$ is true if ψ holds somewhere in the past in a time instant belonging to the interval $[t - t_2, t - t_1]$ and φ is true from then up to now.

Tables 1 and 2 show the MTL semantics of the Timed Declare constraints used in this paper and their graphical notation. Consider, for example, the *response* constraint $\mathbf{G}(a \rightarrow \mathbf{F}_{[t_1,t_2]}b)$. This constraint indicates that if a occurs at a time instant t, b must eventually follow at a time instant belonging to the interval $[t+t_1, t+t_2]$. Note that negative constraints are used to model latency constraints, i.e., constraints stating that at least a minimum amount of time is needed to accomplish a certain operation [12]. For example, the *not response* constraint $\mathbf{G}(a \rightarrow \neg(\mathbf{F}_{[0,t_1]}b))$ indicates that if a occurs at a time instant t, it takes at least t_1 time units to execute b (b cannot be executed in the time interval $[t, t + t_1]$).

2.3 Event Correlations

We use the term event correlation to indicate a mechanism to link two events in a trace. Correlations are defined over event attributes and linked through *relationship operators* between them. For example, two events are correlated if they act upon common data elements of the process or if they are executed by the same resource etc. For a categorization of correlations we refer to [3] in which the following types of correlations are defined: (i) *Property-based correlation*, i.e., events are classified based on a function operating on their attributes. For example, all claim applications referring to an amount greater than 1000 euros are grouped together; (ii) *Reference-based correlation*, i.e., two events are correlated if an attribute of the first event (identifier attribute) and an attribute of the second event (reference attribute) have the same value; (iii) *Moving time-window correlation*, i.e., two events are correlated if they occur within a given duration of one another (e.g., one hour).

In this paper, we use a variation of reference-based correlation according to which two events are correlated if there is a function connecting an attribute of the first event with an attribute of the second event. This function can include operators such as *greater than*, *less than*, *equal to*, and *not equal to*. For example, an event of producing a document is correlated to an event of checking it if the resource that produces the document is different from the one that checks it.

To correctly associate an activation and a target of a given constraint, correlations can be provided by a domain expert or, alternatively, they can be automatically discovered from event logs (as presented in [4]). For discovering event correlations from a log, in [4], the authors first generate all feasible correlations for the constraint under examination, i.e., correlations between comparable attributes of activations and targets of that constraint (comparable attributes are attributes having the same data type). The "goodness" of a (feasible) correlation is then evaluated based on its support.

In this paper, we define the support of a correlation (for positive relation constraints) as the ratio between the number of activations that, through the given correlation, can be linked to *at least* one target and the total number of activations. For negative relation constraints, we define the support of a correlation as the ratio between the number of activations for which there is *no* correlated target and the total number of activations. In this way, for negative relation constraints, the support of a correlation is higher if it allows us to decouple activations and possible targets instead of connecting them. Note that, if we define the support of a correlation in this way, the support of a correlation evaluated for a negative relation constraint can be derived from the support of the same correlation evaluated for the corresponding positive relation constraint. For example, if $support_{response(a,b)}(corr)$ is the support in a log of a correlation *corr* for a response constraint between activities a and b, we have that $support_{not\ response(a,b)}(corr) = 1 - support_{response(a,b)}(corr)$.

3 Discovering Timed Declare Constraints from Logs

In this section, we describe the proposed algorithms for the discovery of Timed Declare constraints from event logs. Each of them can be used to discover constraints referring to different Declare templates. Every algorithm takes as input a log to be processed. Moreover, the presented algorithms work on a given candidate constraint (without time information: the time intervals to be associated to each candidate are determined in a later stage). Therefore, they assume that a list of candidate constraints has been created beforehand. The list of candidate constraints can include all the possible instantiations of a template with each possible combination of event names in the given log (see [11]). This approach can be optimized and the number of candidate constraints can be reduced by using a seminal Apriori algorithm as presented in [10]. Every algorithm requires as input also a correlation to link each activation of the candidate constraint under examination with the corresponding target.

The algorithms produce as output the percentage of traces in the log in which each candidate constraint is non-vacuously satisfied (the support of the candidate constraint in the log) and a vector containing the time distances between each activation and the corresponding target. The support value is used to filter out candidate constraints that are non-vacuously satisfied in a low percentage of log traces (with respect to a user-defined threshold). The time distances are used to identify the time interval $[t_1, t_2]$ characterizing the discovered metric temporal constraint (see Section 3.4).

In the following sections, we describe the algorithms we use for the discovery of positive relation constraints. The same algorithms can be used also for the discovery of negative relation constraints taking into account that, as explained in Section 3.4, the mechanisms for the definition of the time intervals are different in the two cases.

3.1 Timed Response, Precedence, and Responded Existence

The algorithm presented in this section (Algorithm 1) can be used for the discovery of *timed response* constraints. A similar algorithm (with small modifications) is able to discover *timed precedence* and *timed responded existence* constraints.

Algorithm 1. Discovery algorithm for timed response.

```
   Input: log, the event log to be processed; (a, t), activation and target of a candidate constraint; corr a
          selected correlation for linking each activation with the corresponding target
1  define vector timeDistances containing the time distances between each activation and the corresponding
   targets
2  validTracesNumber := 0
3  foreach trace ∈ log do
4      define vector pendingActivations_trace containing events corresponding to pending activations in trace
       trace
5      activated := false
6      foreach event ∈ trace do
7          if event.name = a then
8              activated := true
9              pendingActivations_trace.add(event)
10         if event.name = t then
11             foreach p ∈ pendingActivations_trace do
12                 if isValid(corr, p, event) then
13                     pendingActivations_trace.remove(p)
14                     timeDistances.add(|p.timestamp − event.timestamp|)
15     if pendingActivations_trace.size = 0 && activated then
16         validTracesNumber + +
17 support := validTracesNumber/log.size
   Output: support, the support of the candidate constraint; timeDistances
```

The inputs of this algorithm are an event log *log*, a candidate constraint with activation a and target t and a feasible correlation *corr*. The outputs of the algorithm are vector *timeDistances* containing the time distances between each activation of the candidate constraint and the corresponding targets, and the support of the candidate constraint *support*, i.e., the ratio between the number of traces in which the constraint is non-vacuously satisfied ($validTracesNumber$) and the total number of traces in the event log.

For each trace *trace* in *log*, $pendingActivations_{trace}$ is a vector containing the pending activations in *trace* for the candidate constraint under examination, i.e., the activations that do not have a corresponding target. For each event *event* in *trace*, *event* is a pending activation if its event name is equal to a (*event* is an activation and is not associated to any target yet). In this case *event* is added to $pendingActivations_{trace}$ and the constraint is activated in *trace* (lines 8 and 9). On the other hand, if the event name of *event* is equal to t, *event* is a possible target corresponding to one of the pending activations in $pendingActivations_{trace}$. In this case, if a pending activation p in $pendingActivations_{trace}$ can be correlated to *event* through correlation *corr* (the algorithm checks if this is the case through function $isValid$), then p is removed from the set of pending activations and the time distance between p and the corresponding target *event* is added to *timeDistances* (lines 13 and 14). The candidate constraint under examination is non-vacuously satisfied in *trace*, if, when all the events in *trace* have been processed, the boolean variable *activated* is true (*trace* contains at least one activation of the candidate constraint) and all the activations in *trace* have a corresponding target (i.e., if $pendingActivations_{trace}$ is empty).

The algorithm for the discovery of *timed precedence* constraints is similar to the one described for timed response. The difference is that, for timed precedence constraints, we iterate each trace in the event log (line 6) from the last event to the first one. In the algorithm for the discovery of *timed responded existence*

Algorithm 2. Discovery algorithm for timed alternate response.

```
Input: log, the event log to be processed; (a, t), activation and target of a candidate constraint; corr a
       selected correlation for linking each activation with the corresponding target
define vector timeDistances containing the time distances between each activation and the corresponding
targets
validTracesNumber := 0
foreach trace ∈ log do
    define vector pendingActivations_trace containing events corresponding to pending activations in trace
    trace
    define vector possibleTargets_trace containing possible target events corresponding to a pending
    activation
    violated := false
    activated := false
    foreach event ∈ trace do
        if event.name = a then
            activated := true
            if possibleTargets_trace.size ⩾ 1 && pendingActivations_trace.size = 1 then
                targetFound := false
                previousAct := element ∈ pendingActivations_trace
                foreach p ∈ possibleTargets_trace do
                    if isValid(corr, previousAct, p) then
                        targetFound := true
                        timeDistances.add(|previousAct.timestamp − p.timestamp|)

                if !targetFound then
                    violated := true

            if possibleTargets_trace.size = 0 && pendingActivations_trace.size = 1 then
                violated := true

            if possibleTargets_trace.size ⩾ 1 then
                pendingActivations_trace.removeAll()

            possibleTargets_trace.removeAll()
            pendingActivations_trace.add(event)

        if event.name = t then
            possibleTargets_trace.add(event)

    if possibleTargets_trace.size ⩾ 1 && pendingActivations_trace.size = 1 then
        targetFound := false
        previousAct := element ∈ pendingActivations_trace
        foreach p ∈ possibleTargets_trace do
            if isValid(corr, previousAct, p) then
                targetFound := true
                timeDistances.add(|previousAct.timestamp − p.timestamp|)

        if !targetFound then
            violated := true

        pendingActivations_trace.removeAll()

    if pendingActivations_trace.size = 0 && activated && !violated then
        validTracesNumber + +

support := validTracesNumber/log.size
Output: support, the support of the candidate constraint; timeDistances
```

constraints, we iterate each trace in the log in both directions so that each activation can have zero, one, or two possible targets. If there are no possible target, there is a violation in the trace. If there is only one possible target we evaluate the time distance between the activation and the corresponding target. If there are two possible targets, we consider the one with the smallest time distance from the considered activation.

3.2 Timed Alternate Response and Alternate Precedence

The algorithm presented in this section (Algorithm 2) can be used for the discovery of *timed alternate response* constraints. In the same way as illustrated for timed response and timed precedence constraints, also in this case a similar algorithm is able to discover *timed alternate precedence* constraints.

The inputs of this algorithm are an event log *log*, a candidate constraint with activation a and target t and a feasible correlation *corr*. The outputs of the algorithm are vector *timeDistances* containing the time distances between each activation of the candidate constraint and the corresponding targets, and the support of the candidate constraint *support*.

For each trace *trace* in *log*, $pendingActivations_{trace}$ is a vector containing the pending activations in *trace* for the candidate constraint under examination. $possibleTargets_{trace}$ is a vector containing the possible targets for the last pending activation encountered in *trace* and added to $pendingActivations_{trace}$. For each event *event* in *trace*, if the event name of *event* is equal to a, the constraint is activated in *trace* (line 10). In this case, the algorithm checks if there is one pending activation (*previousAct*) in $pendingActivations_{trace}$ and there is at least one possible corresponding target in $possibleTargets_{trace}$ for this activation (line 11). If this is the case, the algorithm checks if *previousAct* can be correlated to one of the possible targets in $possibleTargets_{trace}$ (through function $isValid$). If *previousAct* has a correlated target, then the time distance between *previousAct* and the corresponding target is added to *timeDistances* (line 17), otherwise a violation is detected (line 19).

If *event* is an activation, there is already one pending activation in $pendingActivations_{trace}$ and there are no possible corresponding targets for this activation ($possibleTargets_{trace}$ is empty), then a violation is detected (line 21) since an alternate response constraint does not allow two activations to occur one after another without any target in between (see Table 1).

Finally, if *event* is an activation and *event* is preceded by at least one possible target, then all the elements in $pendingActivations_{trace}$ are deleted (line 23). In all cases, when *event* is an activation, all elements in $possibleTargets_{trace}$ are deleted and *event* is added to $pendingActivations_{trace}$ (lines 24 and 25). If *event* is a target of the candidate constraint under examination, *event* is added to $possibleTargets_{trace}$ (line 27).

After that the last event in *trace* has been processed, the algorithm checks again if there is still one pending activation in $pendingActivations_{trace}$ with at least one possible corresponding target in $possibleTargets_{trace}$. If this is the case, the algorithm checks if the pending activation can be correlated to one of the possible targets. The candidate constraint is non-vacuously satisfied in *trace*, if, when all the events in *trace* have been processed, the boolean variable *activated* is true (*trace* contains at least one activation of the candidate constraint), $pendingActivations_{trace}$ is empty and the boolean variable *violated* is false.

3.3 Timed Chain Response and Chain Precedence

The algorithm presented in this section (Algorithm 3) can be used for the discovery of *timed chain response* constraints. A similar algorithm is able to discover *timed chain precedence* constraints.

The inputs of this algorithm are an event log *log*, a candidate constraint with activation a and target t and a feasible correlation *corr*. The outputs of the

Algorithm 3. Discovery algorithm for timed chain response.

```
   Input: log, the event log to be processed; (a, t), activation and target of a candidate constraint; corr a
          selected correlation for linking each activation with the corresponding target
1  define vector timeDistances containing the time distances between each activation and the corresponding
   targets
2  validTracesNumber := 0
3  foreach trace ∈ log do
4      define vector pendingActivations_trace containing events corresponding to pending activations in trace
       trace
5      violated := false
6      activated := false
7      foreach event ∈ trace do
8          if pendingActivations_trace.size = 1 then
9              p := element ∈ pendingActivations_trace
10             if event.name = t then
11                 if isValid(corr, p, event) then
12                     timeDistances.add(|p.timestamp − event.timestamp|)
13                 else
14                     violated := true
15             else
16                 violated := true
17             pendingActivations_trace.removeAll()
18         if event.name = a then
19             pendingActivations_trace.add(event)
20             activated := true
21     if pendingActivations_trace.size = 0 && activated && !violated then
22         validTracesNumber + +
23 support := validTracesNumber/log.size
   Output: support, the support of the candidate constraint; timeDistances
```

algorithm are vector *timeDistances* containing the time distances between each activation of the candidate constraint and the corresponding targets, and the support of the candidate constraint *support*.

For each trace *trace* in *log*, $pendingActivations_{trace}$ is a vector containing the pending activations in *trace* for the candidate constraint under examination. For each event *event* in *trace*, if *event* is an activation, it is added to $pendingActivations_{trace}$ and the constraint is activated in *trace* (lines 19 and 20). If $pendingActivations_{trace}$ contains an element, and *event* is a target event, then the algorithm checks if the pending activation in $pendingActivations_{trace}$ can be correlated to *event* through correlation *corr* (the algorithm checks if this is the case through function $isValid$). If the two events can be correlated, the time distance between the pending activation and the corresponding target *event* is added to *timeDistances*, otherwise a violation is detected (line 14). A violation is detected also if $pendingActivations_{trace}$ contains an activation and the current event *event* is not a target event (line 16) since according to the semantics of the chain response, a target should immediately follow an activation. The candidate constraint is non-vacuously satisfied in *trace*, if, when all the events in *trace* have been processed, the boolean variable *activated* is true, $pendingActivations_{trace}$ is empty and the boolean variable *violated* is false.

3.4 Time Intervals Identification

In this section, we describe how the time distances generated by the discovery algorithm described so far can be used to identify the time interval $[t_1, t_2]$ characterizing the discovered metric temporal constraint. Suppose that d_{min} and d_{max}

are the minimum and the maximum time distances between an activation and the corresponding target of the candidate constraint under examination. Then, if the candidate constraint is a positive relation constraint, we simply derive that $t_1 = d_{min}$ and $t_2 = d_{max}$.

Also in the case of a negative relation constraint, if an activation occurs at time t all the possible targets for that activations occur at a time point belonging to the interval $[t+d_{min}, t+d_{max}]$. However, in case of a negative relation constraint, the constraint is valid outside this interval. Therefore, we can associate two intervals to the candidate constraint (thus discovering two constraints from it), i.e., $[0, d_{min}[$ and $]d_{max}, \infty[$.

To deal with cases in which logs contain noise, it is also possible to choose the boundaries of the time intervals by removing outliers from the discovered set of time distances. This can easily be done by removing time distances that deviate more than a given threshold from the average.

4 Validation

The discovery algorithms presented in this paper have been implemented in a ProM plug-in. The plug-in takes an event log and a set of correlations (one for each candidate constraint) as input and produces a Declare map consisting of a set of Timed Declare constraints. We have conducted our validation by applying the implemented plug-in to the real-life event logs provided for the BPI challenges 2011 and 2012. In Sections 4.1 and 4.2, we present the two case studies.

4.1 A Case Study Based on the BPI Challenge 2012

The first case study we discuss is based on the application of the proposed approach to the event log provided for the BPI challenge 2012 [8] and taken from a Dutch financial institute. The event log pertains to an application process for personal loans or overdrafts. It contains 262,200 events distributed across 36 event classes (i.e., each event can be associated to one of 36 different possible event names) and includes 13,087 cases. The amount requested by the customer is indicated in the case attribute *AMOUNT_REQ*. In addition, the log contains the standard XES attributes for events: *concept:name*, *lifecycle:transition*, *time:timestamp*, and *org:resource*.[3] The event log merges three intertwined sub processes. Therefore, in each case, events belonging to different sub processes can occur. These events should be correlated with each other.

Fig. 1 shows an excerpt of the table that ProM generates and that contains the correlations discovered from the log for each candidate constraint. In the table, for each candidate constraint (identified through its Declare template, its activation A and its target T), a list of correlations is specified. For each correlation, its support and degree of disambiguation is indicated. As well as the correlation support, the *degree of disambiguation* is a metric that helps the user

[3] XES (eXtensible Event Stream) is an XML-based standard for event logs proposed by the IEEE Task Force on Process Mining (`www.xes-standard.org`).

in understanding how good a correlation is. In particular, for each correlation, its degree of disambiguation is defined as the ratio between the number of ambiguous activations that can be disambiguated with the considered correlation and the total number of ambiguous activations.

In this example, for each candidate constraint, we choose the correlation "A.org:resource = T.org:resource" because, in all cases, this correlation has a high support and a good degree of disambiguation. Therefore, we assume that activation and target of a candidate constraint are correlated if and only if the corresponding activities are executed by the same resource. Then, we discovered *timed response constraints* with a minimum support of 30%.[4]

The plug-in allows the user to choose the time granularity to express the boundaries of the discovered time intervals. In a first attempt, we have chosen *days* as time granularity. The discovered map is shown in Fig. 2. From this map it is evident that *days* is not the best time granularity since, in many cases, with this granularity, the boundaries of the discovered time intervals are equal to 0. Therefore, we tried to choose a finer granularity (*hours*). In Fig. 3, we can see the result. The granularity of *hours* is clearly more suitable for this example. In the discovered map, we can see that, for example, the maximum delay between activation *W_Completeren aanvraag-complete* and the corresponding target *W_Nabellen offertes-start* is of 719.52 hours (around 1 month). In some cases, the two activities are executed in a time span smaller than 1 hour (the delay between the two activities can be 0).

4.2 A Case Study Based on the BPI Challenge 2011

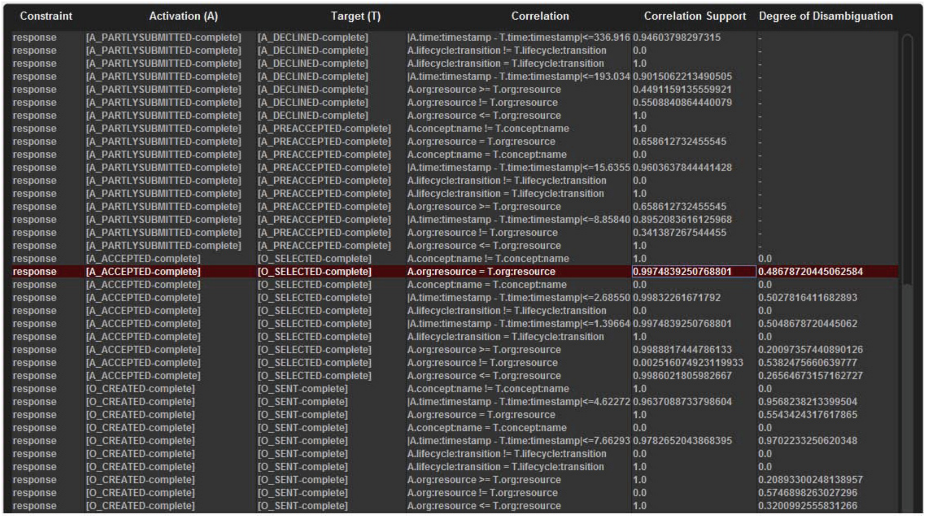
The second case study we present here is based on the application of the proposed approach to the BPI challenge 2011 event log [1] pertaining to the treatment of patients diagnosed with cancer in a large Dutch academic hospital. The event log contains 1,143 cases and 150,291 events distributed across 623 event classes. Each case in this event log is related to a different patient. We suppose, in this case, that all the events in a single case are correlated to each other and, for this reason, we adopt a conservative approach in the discovery of Timed Declare constraints. Our plug-in supports this when the user does not specify any correlation for a candidate constraint.

We discovered from this log *timed responded existence constraints* with a minimum support of 60%. The discovered map is shown in Fig. 4. The time granularity chosen for this map is *days*. The map shows that, for example, the maximum delay between activation *administrative fee - the first pol* and the corresponding target *First outpatient consultation* is of 1,127 days (more than 3 years). In some cases, the two activities are executed in the same day (the delay can be 0).

5 Conclusion

The recent contributions in the context of declarative process discovery do not take the time dimension into consideration. In this paper, we present algorithms

[4] The selected support threshold is low because the average support of the timed response constraints discovered from this log is low.



Constraint	Activation (A)	Target (T)	Correlation	Correlation Support	Degree of Disambiguation
response	[A_PARTLYSUBMITTED-complete]	[A_DECLINED-complete]	|A.time:timestamp - T.time:timestamp|<=336.916	0.94603798297315	-
response	[A_PARTLYSUBMITTED-complete]	[A_DECLINED-complete]	A.lifecycle:transition != T.lifecycle:transition	0.0	-
response	[A_PARTLYSUBMITTED-complete]	[A_DECLINED-complete]	A.lifecycle:transition = T.lifecycle:transition	1.0	-
response	[A_PARTLYSUBMITTED-complete]	[A_DECLINED-complete]	|A.time:timestamp - T.time:timestamp|<=193.034	0.9015062213490505	-
response	[A_PARTLYSUBMITTED-complete]	[A_DECLINED-complete]	A.org:resource >= T.org:resource	0.4491159135559921	-
response	[A_PARTLYSUBMITTED-complete]	[A_DECLINED-complete]	A.org:resource != T.org:resource	0.5508840864440079	-
response	[A_PARTLYSUBMITTED-complete]	[A_DECLINED-complete]	A.org:resource <= T.org:resource	1.0	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	A.concept:name != T.concept:name	1.0	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	A.org:resource = T.org:resource	0.658612732455545	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	A.concept:name = T.concept:name	0.0	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	|A.time:timestamp - T.time:timestamp|<=15.6355	0.9603637844441428	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	A.lifecycle:transition != T.lifecycle:transition	0.0	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	A.lifecycle:transition = T.lifecycle:transition	1.0	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	A.org:resource >= T.org:resource	0.658612732455545	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	|A.time:timestamp - T.time:timestamp|<=8.85840	0.8952083616125968	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	A.org:resource != T.org:resource	0.341387267544455	-
response	[A_PARTLYSUBMITTED-complete]	[A_PREACCEPTED-complete]	A.org:resource <= T.org:resource	1.0	-
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	A.concept:name != T.concept:name	1.0	0.0
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	A.org:resource = T.org:resource	0.9974839250768801	0.48678720445062584
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	A.concept:name = T.concept:name	0.0	0.0
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	|A.time:timestamp - T.time:timestamp|<=2.68550	0.99832261671792	0.5027816411682893
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	A.lifecycle:transition != T.lifecycle:transition	0.0	0.0
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	|A.time:timestamp - T.time:timestamp|<=1.39664	0.9974839250768801	0.5048678720445062
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	A.lifecycle:transition = T.lifecycle:transition	1.0	0.0
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	A.org:resource >= T.org:resource	0.9988817444786133	0.20097357440890126
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	A.org:resource != T.org:resource	0.002516074923119933	0.5382475660639777
response	[A_ACCEPTED-complete]	[O_SELECTED-complete]	A.org:resource <= T.org:resource	0.9986021805982667	0.26564673157162727
response	[O_CREATED-complete]	[O_SENT-complete]	A.concept:name != T.concept:name	1.0	0.0
response	[O_CREATED-complete]	[O_SENT-complete]	|A.time:timestamp - T.time:timestamp|<=4.62272	0.9637088733798604	0.9568238213399504
response	[O_CREATED-complete]	[O_SENT-complete]	A.org:resource = T.org:resource	1.0	0.5543424317617865
response	[O_CREATED-complete]	[O_SENT-complete]	A.concept:name = T.concept:name	0.0	0.0
response	[O_CREATED-complete]	[O_SENT-complete]	|A.time:timestamp - T.time:timestamp|<=7.66293	0.9782652043868395	0.9702233250620348
response	[O_CREATED-complete]	[O_SENT-complete]	A.lifecycle:transition != T.lifecycle:transition	0.0	0.0
response	[O_CREATED-complete]	[O_SENT-complete]	A.lifecycle:transition = T.lifecycle:transition	1.0	0.0
response	[O_CREATED-complete]	[O_SENT-complete]	A.org:resource >= T.org:resource	1.0	0.20893300248138957
response	[O_CREATED-complete]	[O_SENT-complete]	A.org:resource != T.org:resource	0.0	0.5746898263027296
response	[O_CREATED-complete]	[O_SENT-complete]	A.org:resource <= T.org:resource	1.0	0.3200992555831266

Fig. 1. Correlations selection

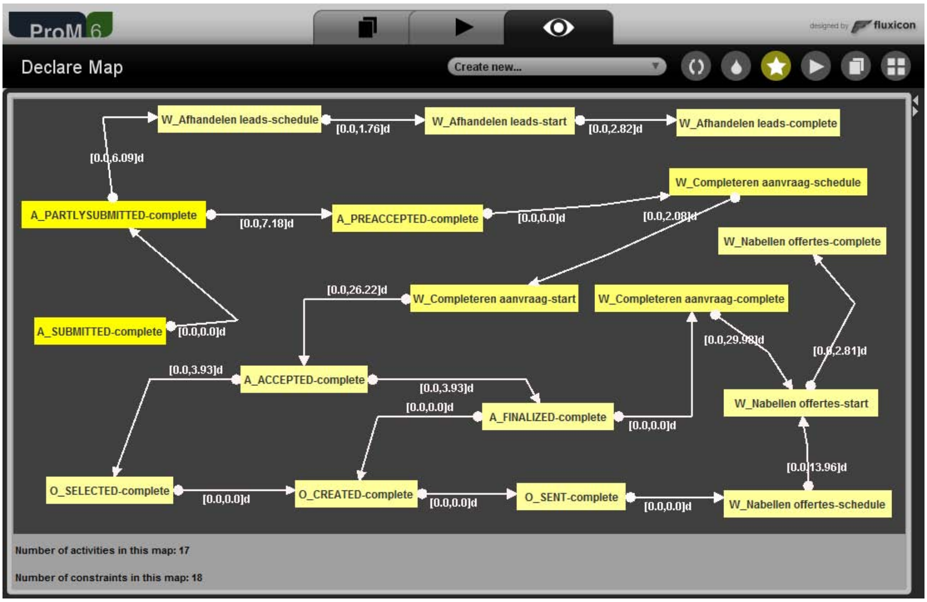


Fig. 2. Discovered *timed response constraints* from the log provided for the BPI Challenge 2012 (time granularity: *days*)

for discovering Timed Declare constraints from event logs. We use correlations to connect each activation of a constraint with the corresponding target thus improving the reliability of the discovered metric temporal constraints. The proposed algorithms can also be used for discovering negative relation constraints.

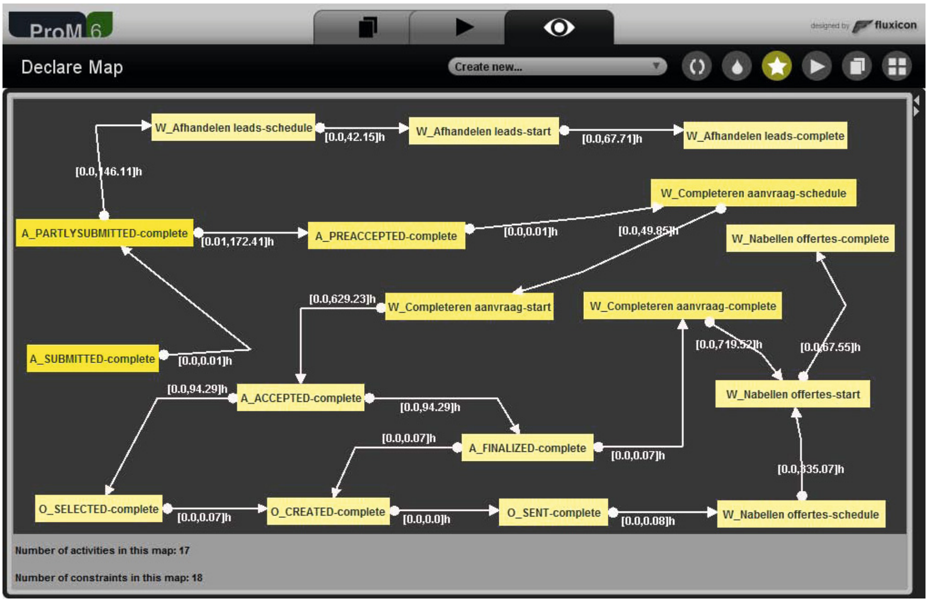


Fig. 3. Discovered *timed response constraints* from the log provided for the BPI Challenge 2012 (time granularity: *hours*)

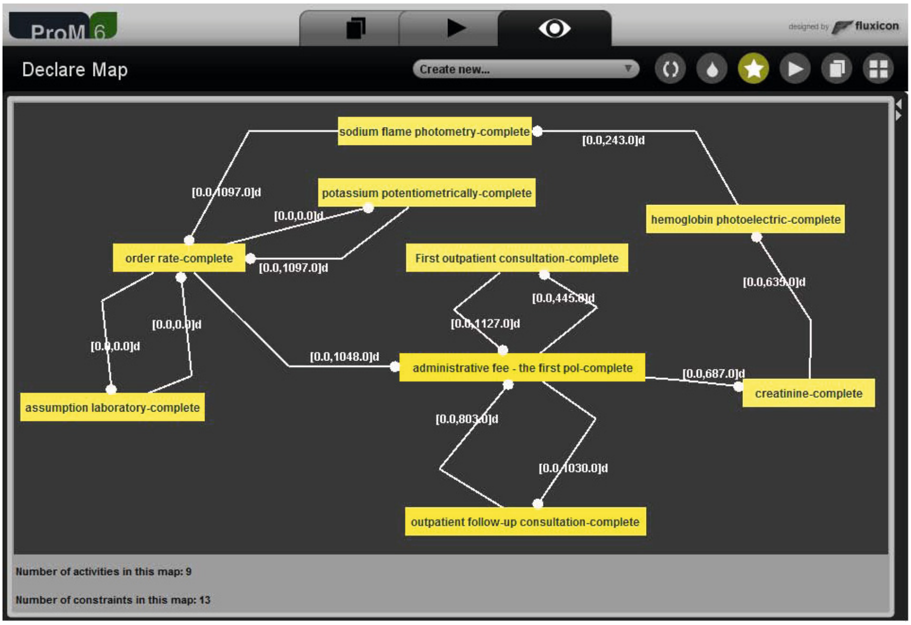


Fig. 4. Discovered *timed responded existence constraints* from the log provided for the BPI Challenge 2011 (time granularity: *days*)

To this aim, we have extended the notion of correlation support also to cover this group of constraints. Our evaluation using real-life logs demonstrates that the proposed approach is applicable in real-life settings.

References

1. 3TU Data Center: BPI Challenge 2011 Event Log (2011), doi:10.4121/uuid:d9769f3d-0ab0-4fb8-803b-0d1120ffcf54
2. van der Aalst, W., Pesic, M., Schonenberg, H.: Declarative Workflows: Balancing Between Flexibility and Support. Computer Science - R&D, 99–113 (2009)
3. Barros, A., Decker, G., Dumas, M., Weber, F.: Correlation Patterns in Service-Oriented Architectures. In: Dwyer, M.B., Lopes, A. (eds.) FASE 2007. LNCS, vol. 4422, pp. 245–259. Springer, Heidelberg (2007)
4. Bose, R.P.J.C., Maggi, F.M., van der Aalst, W.M.P.: Enhancing declare maps based on event correlations. In: Daniel, F., Wang, J., Weber, B. (eds.) BPM 2013. LNCS, vol. 8094, pp. 97–112. Springer, Heidelberg (2013)
5. Burattin, A., Maggi, F.M., van der Aalst, W.M.P., Sperduti, A.: Techniques for a Posteriori Analysis of Declarative Processes. In: EDOC, pp. 41–50 (2012)
6. Chesani, F., Lamma, E., Mello, P., Montali, M., Riguzzi, F., Storari, S.: Exploiting Inductive Logic Programming Techniques for Declarative Process Mining. In: Jensen, K., van der Aalst, W.M.P. (eds.) ToPNoC II. LNCS, vol. 5460, pp. 278–295. Springer, Heidelberg (2009)
7. Di Ciccio, C., Mecella, M.: Mining constraints for artful processes. In: Abramowicz, W., Kriksciuniene, D., Sakalauskas, V. (eds.) BIS 2012. LNBIP, vol. 117, pp. 11–23. Springer, Heidelberg (2012)
8. van Dongen, B.: Bpi challenge 2012 (2012), `http://dx.doi.org/10.4121/uuid:3926db30-f712-4394-aebc-75976070e91f`
9. Kupferman, O., Vardi, M.Y.: Vacuity Detection in Temporal Model Checking. International Journal on Software Tools for Technology Transfer, 224–233 (2003)
10. Maggi, F.M., Bose, R.P.J.C., van der Aalst, W.M.P.: Efficient Discovery of Understandable Declarative Models from Event Logs. In: Ralyté, J., Franch, X., Brinkkemper, S., Wrycza, S. (eds.) CAiSE 2012. LNCS, vol. 7328, pp. 270–285. Springer, Heidelberg (2012)
11. Maggi, F.M., Mooij, A.J., van der Aalst, W.M.P.: User-Guided Discovery of Declarative Process Models. In: CIDM, pp. 192–199 (2011)
12. Montali, M.: Specification and Verification of Declarative Open Interaction Models: a Logic-Based Approach. Ph.D. thesis, University of Bologna (2009)
13. Pichler, P., Weber, B., Zugal, S., Pinggera, J., Mendling, J., Reijers, H.A.: Imperative Versus Declarative Process Modeling Languages: An Empirical Investigation. In: Daniel, F., Barkaoui, K., Dustdar, S. (eds.) BPM Workshops 2011, Part I. LNBIP, vol. 99, pp. 383–394. Springer, Heidelberg (2012)
14. Reijers, H.A., Slaats, T., Stahl, C.: Declarative modeling-an academic dream or the future for bpm? In: Daniel, F., Wang, J., Weber, B. (eds.) BPM 2013. LNCS, vol. 8094, pp. 307–322. Springer, Heidelberg (2013)
15. Thati, P., Roşu, G.: Monitoring algorithms for metric temporal logic specifications. Electron. Notes Theor. Comput. Sci. 113, 145–162 (2005)
16. Westergaard, M., Maggi, F.M.: Looking into the future: Using timed automata to provide a priori advice about timed declarative process models. In: Meersman, R., et al. (eds.) OTM 2012, Part I. LNCS, vol. 7565, pp. 250–267. Springer, Heidelberg (2012)

Generating a Business Model Canvas through Elicitation of Business Goals and Rules from Process-Level Use Cases

Carlos E. Salgado[1], Juliana Teixeira[1], Ricardo J. Machado[1], and Rita S.P. Maciel[2]

[1] Universidade do Minho, Guimarães, Portugal
[2] Universidade Federal da Bahia, Salvador, Brasil
carlos.salgado@algoritmi.uminho.pt,
juliana.teixeira@research.ccg.pt, rmac@dsi.uminho.pt,
ritasuzana@dcc.ufba.br

Abstract. Business Models play a pivotal role in organizations, especially in building bridges and enabling the dialogue between business and technological worlds. Complementarily, while Use Cases are one of the most popular techniques for eliciting requirements in the design of Information Systems, Business Goals and Business Rules associate with Business Process Use Cases to compose a Business Model base structure. However, methods for relating Business Processes, Goals and Rules (PGR) are scarce, dissonant or highly analyst-dependent. In this sense, we propose a two-step method to help in guiding the elicitation of Business Goals and Rules from Process-level Use Cases, and their mapping to a Business Model representation. As a result, a solution Business Model generated by aligning the resulting trios (PGR) with a Business Model Canvas is presented to the organization stakeholders for review, validation and further negotiation.

Keywords: Business Model, Business Goals, Business Rules, Business Use Cases, Requirements Elicitation.

1 Introduction

Business Models play an ever more pivotal role in the development and continued management of Information Systems (IS). Nevertheless, recent literature review on Business Models (BM) results show that there is no agreement on what a BM is, although some emerging common themes already exist [1]. The BM artifact, as a conceptual tool that contains a set of elements and their relationships, expressing the business logic of a specific firm and the value it offers, is seen as crucial for improving the Business-IS/IT dialogue. Although it has had substantial attention from both academics and practitioners, its growing associated literature is still young and dispersed.

Our recent work in generating a BM in ill-defined contexts, within a RUP-based approach and grounded on reference model representations, stands as a contribution

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 276–289, 2014.

inside this topic [2]. The use of Process-level Use Cases, together with Business Goals and Rules associated information (PGR), allows developing an activity direct-mapped BM to present to stakeholders for validation. The use and adaptation of 'standard' methods and techniques to infer goals and rules requirements from scenarios and process-like diagrams, mapping backwardly the traditional business to process workflow, could also allow for better and continuous alignment between Business and IS/IT, with improved traceability.

Accordingly, the knowledge represented in terms of goals, rules and methods can make reengineering tasks more systematic and effective [3]. Whether it involves the development of a new system or the reengineering of business processes, decisions about what goals to pursue and on selecting the appropriate strategies to achieve them are always vital. The discovery of goals and rules is part of requirements elicitation, recognized as one of the most critical activities of software development, with many prescribed methods and techniques.

However, it is virtually impossible to define a unified model for the elicitation process, due to the constantly changing needs associated to requirements activities. Even if specific methodologies, broken down into multiple steps, describe general approaches and overall principles to assist analysts in understanding needs, only the experienced analyst understands intuitively which method or technique is effective, in each circumstance, and is able to apply it [4]. This raises issues as lack of formality, analyst dependency and difficulties for less experienced analysts, all added to the Business-IS/IT dialogue gap.

Our approach tries to obviate to these, presenting a method, to guide the analyst in the elicitation of Business Goals and Rules from Process-level Use Cases, and transforming them, in order to arrive at a BM representation. This later can then be presented to the involved stakeholders for review, validation and further negotiation. As the entire method follows a model-based approach, the changes agreed upon could be traced back to the original Use Cases, allowing for requirements traceability and a Business-IS/IT aligned solution.

This document follows with a background research review on Business Model representations and on diverse methods, techniques and guidelines for the elicitation of Business Goals and Rules. Then, we present our proposed two-step method, covering the elicitation of Business Goals and Rules from Process-level Use Cases, and their mapping to a BM representation, resulting in a generated BM aligning our PGR trios with the original Use Cases. Next we apply and demonstrate it in a live project setting. Consequently, an analysis of the results obtained and future work directions are discussed in section 5. Finally, some conclusions are drawn for this paper.

2 Related Research

This section presents related research regarding Business Model representations, and Business Goals and Rules elicitation approaches. For the BM topic, it focuses solely on the Business Model Canvas (BMC) [5] and its early connection with the Balanced

Scorecard (BSC) [6], mainly due to their popularity in Business-IS/IT communities. Relating to methods and techniques for eliciting goals and rules, it falls in their associated combination of checklists and guidelines from Rational Unified Process (RUP) [7], and in the business plans representation of Business Motivation Model (BMM) [8]. Notwithstanding other elicitation methods and techniques as i* [9] or KAOS [10], this choice is due to the more complete and business oriented side of RUP and BMM, which help in defining the business requirements specification for business modeling, and promote the Business and IS/IT alignment questions that are comprised in process-oriented approaches.

2.1 Business Models

The BMC, a strategic management template for developing new or documenting existing business models, currently stands out as one of the preferred tools for their generation, especially in business related audiences. The BMC is based on the Business Model Ontology proposed by Osterwalder [5], where the formal descriptions of the business become the building blocks for its activities. These are divided in nine different business conceptualizations, organized by four dimensions: Infrastructure, Offering, Customers and Finance. In turn, this division was based on the early work of Kaplan and Norton [6] with the BSC four perspectives: Financial, Customer, Internal Business Process and Learning & Growth.

Regarding BMC, the value proposition describes which customers' problems are solved and why the offer is more valuable than the competitors' similar ones. Customers themselves are analyzed in segments and relationships, while the distribution channels means to illustrate how the customer needs to be reached and by whom. Resources are transformed through the key activities into the final product or service, which depends on external partner networks. Financial information focus on cost and revenue, where cost structures should be aligned to the core ideas and revenue streams must mirror the value assigned by customers, in terms of how much they are willing to pay, and how they will perform transactions.

Accordingly, BSC management perspectives track its four high-level types of measures, using this approach to identify corporate objectives within each of the categories and seeking to align the balanced set of measures with the organization strategy. It uses a top-down method that defines the organization strategic goals, passing these goals downward, whose success results form a strategy-focused organization, derived from strategy maps and balanced scorecards. A strategy map shows how the organization creates value and describes how organization key business objectives align within the four BSC perspectives to support corporate strategies.

BMC and BSC are two different but complementary tools to achieve innovation, tactical directions and action plans in an existing or planned organization. While BMC determines part of the business strategy, BSC is aimed to track implementation and ensure that the organization strategy is executed. Recent research by [11], classified BMC and twenty nine other relevant literature sources on business model, with BMC obtaining interesting global results: positive on 66,7% of all the criteria analyzed, checked on all of the top six criteria items and on 50% (six out of twelve) of the second-level ones.

2.2 Business Goals Elicitation

A recurrent question in research over Business Goals (BG) elicitation is that Use Case (UC) notation is intended for functional requirements and not non-functional requirements, which oversimplifies assumptions on the problem domain. Nowadays, in order for a software system to be of value, it should meet both functional and non-functional requirements, these last by using a goal-oriented representation [12]. In recent years, goal-oriented requirements engineering (GORE) current states and trends from the viewpoints of both academia and industry have been fully scrutinized, with results pointing for goal models to be useful for supporting the decision making process in the early requirements phase [13].

GORE is generally complementary to other approaches, well suited to analyzing requirements early in the software development cycle, especially with respect to non-functional requirements. Goal models represent an essential tool for requirements engineers, system architects, and software developers, but their analysis and evaluation also lead to many challenges [14]. A great variety of techniques for analyzing goal models have been proposed in recent years, but, on the other hand, this diversity creates a barrier for widespread adoption of such techniques, also due to the lack of guidance in literature on which one to choose [15].

According to the RUP guidelines, the characteristics of a good business use-case model include the alignment between UC and strategy, with at least one BG for each UC. This is supported by the BSC classification technique, which allows clarifying the strategic objectives and identifying the critical areas of the organization. Its purpose is to translate the strategy into goals at different levels, and to provide concrete, measurable objectives, which can be directly supported by business processes, ensuring that these are aligned with the business strategy. A BG is a requirement that must be satisfied by the business and directly connected with the UC model, the "glue" between business strategy and business UC.

In BMM, the fundamental idea is to develop a BM through the definition of Ends and Means of business plans. Among the Ends are the goals and objectives to achieve, which may be either a Vision, an overall image of what the organization wants to be, or Desired Results, more specific intents the organization wants to achieve. These are supported by Courses of Action, Goals supported by Strategies and Objectives achieved by Tactics. A Goal is a statement about a state or condition of the enterprise to be brought about or sustained through appropriate Means (Rules-associated), while an Objective quantifies a Goal, providing the measures to determine whether it is being achieved, consistently with the SMART[1] criteria.

2.3 Business Rules Elicitation

Business Rules (BR) are an important artifact in the requirement elicitation process of IS and a vital part in its development cycle, as they describe ongoing policies, procedures, and constraints, which concern an organization in order to achieve its

[1] SMART: specific, measurable, attainable, relevant and time-based.

business goals and objectives. In recent years, much research effort has been done in order to unlock this valuable asset that many organizations have concealed [16]. Its concept has been examined from different points of view, whether as extensions of business goals, or as limitations or constraints on business activities. By structuring, organizing and expressing tactics and policies in a way that is close to business viewpoints, it helps collecting and organizing supports for the implementation of change on a business level for the associated IS [17].

It is important for software to evolve according to changes in its business environment, having BR as an integral part of the software system, its management and evolution. This improves requirements traceability in design as well as minimizes the efforts of changes, as when requirements are systematically identified and linked to design elements, these are easier to locate and implement [18]. Even so, the quality of software engineering projects suffers due to the large gap between the way stakeholders present their requirements and the way analysts capture and express those requirements, with representation of BR as one of the problems. Another issue detected in the analysis of current IS requirements elicitation techniques is that they tend to be much analyst-oriented and dependant [19].

Regarding the RUP approach, as in Goals elicitation it redirects us to BSC techniques to answer WHY questions. In respect to Rules, it states the HOW vision and the formal vs. natural language issues. While formalism permits automation, where OCL (UML) is a possibility, natural language stands as the stakeholders preferred technique as it allows for a more direct description and understanding of the business. By definition, BR are declarations of policies or conditions, and might apply always (invariants) or under a specific condition, being constructed by constraint (restricting its structure and behavior) or by derivation (inferring facts from other facts). Their purpose is to define specific constraints or invariants that must be satisfied by the business.

In BMM, Rules are represented through Means, which may be either a Mission, a Course of Action (Strategy or Tactic), or a Directive (Business Policy or Rule items) to employ in achieving Ends. Directives govern Courses of Action and support the achievement of Ends, but they cannot stand on their own, while Courses of Action can be formulated based on Directives, serving as its source, or can be defined to support the achievement of a Desired Result. Strategies are implemented by Tactics and usually channel efforts towards Goals, while Tactics towards Objectives. BR items guide the Business Process and govern its Course of Action. A Business Policy provides the basis for BR items and governs Business Processes, existing to control, guide and shape the Strategies and Tactics.

3 Generating a Business Model Canvas

In a previous work [2], we proposed the adaptation of standard techniques to infer goals and rules from scenarios and process-like diagrams, mapping backwardly the traditional business to process workflow, which helped in building a business motivation model and defining a strategy for the information system. With an approach based on a BMM representation and guided by a RUP-based backward transformation from process to business, it could allow for better and continuous alignment between Business and IS/IT, with improved traceability.

Following this research work, supported in the previously proposed PGR metamodel [2], we now aim to propose a method to guide the analyst in the elicitation of goals and rules from use cases, and transforming them, in so generating a business-oriented business model for an IS. This is achieved by combining the use of Business Goals and Rules elicited from Business Process Use Cases in a BSC structure, and then performing their mapping to a BMC panel.

Our proposed method is composed by two steps (Fig. 1). The first step aims to elicit and represent the PGR business-side information by following a 'standard' referential, spanning the four perspectives of the BSC, in so improving the consistency of the use cases coverage. The second step analyses and maps each previous elicited item in an adequate section of the BMC panel, linking them to the more abstract level of business modeling, thus delivering an integrated business model to present to stakeholders.

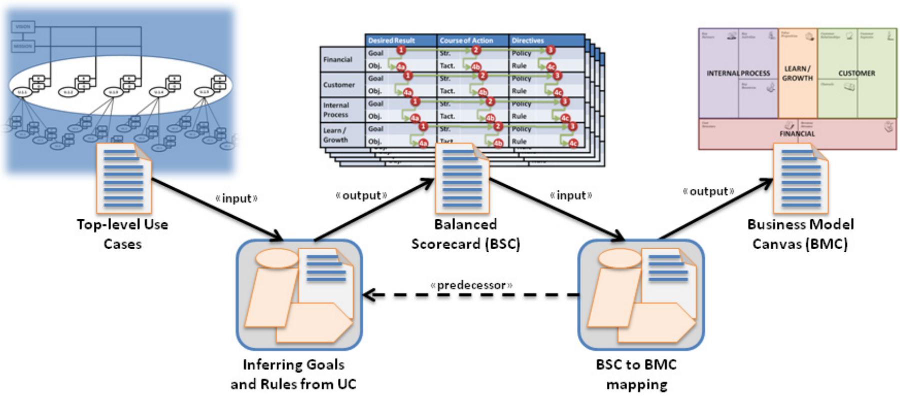


Fig. 1. Process overview of our approach

3.1 Step I – Inferring Goals and Rules from Use Cases

The first step starts with the elicited top-level UC for the proposed IS and involves two iterations, one for each UC and another one for each BSC perspective (Financial, Customer, Internal Business Process and Learn & Growth), with the added BMM representational elements (Goal, Objective, Strategy, Tactic, Business Policy and Business Rule item).

Inside the double-iteration, there are three tasks to be performed (Fig. 2), covering the elicitation of Goals and Rules, with its associated strategies and policies, for each root UC:

1. Envision the UC associated **Goal**;
2. Determine its governing **Strategy**;
3. Associate the controlling **Business Policy**.

These are the more abstract BMM items, the ones preferable to start with due to their business nature, as they should be easier to elicit using the available business documentation. Also, depending on the project, the elicitation of these first elements could be enough for the generation of a high-level, more abstract BM to present to stakeholders for review.

Thereafter, in other situations, a second three tasks sequence (Fig. 2), related to the more concrete and specific goals and rules items, can be performed:

4. a) Define a (SMART) **Objective**, associated to **Goal**;
4. b) Determine a **Tactic**, associated to **Strategy**;
4. c) Delineate a **Business Rule** item, associated to **Business Policy**.

Here, more detailed, concrete information is needed from the project documentation to be able to elicit these items. Although not all fields are mandatory, it is important to fill the most part of them for better system specification and future implementation purposes.

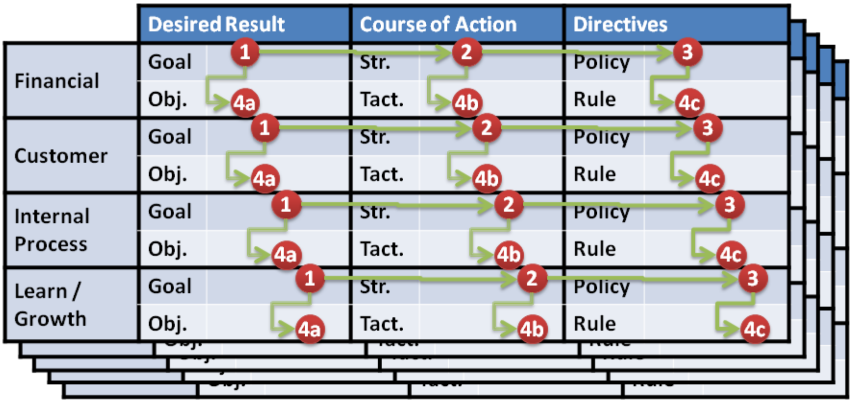


Fig. 2. Tasks for step 1 of our proposed method

All these tasks follow the guidelines and checklists from RUP [7] and BMM [8], as referred in sections 2.2 and 2.3, for the elicitation of Business Goals and Rules, and its associated inner constituent elements. Any further knowledge of other associated techniques by the analysts involved, as well as heuristics associated to their previous experience in the specific domain of the project, are valuable to aid in these tasks.

3.2 Step II – BSC to BMC Mapping

As stated earlier, BMC stands as one of the preferred tools for the generation of business models, especially in business related audiences. Also, BMC relates its roots with BSC, an also popular strategy performance management tool.

According to the Business Model Ontology work [6], the nine elements of the BMC relate directly to the four perspectives of BSC (Fig. 3), namely:

- Financial – Cost Structure and Revenue Streams;
- Customer – Customer Relationships, Channels and Customer Segments;
- Internal Business Process – Key Partners, Activities and Resources;
- Learn & Growth – Value Propositions.

Therefore, our proposal for the mapping of the sentences from our BSC-like structure to the BMC panel follows on this same line of thought: each sentence in BSC maps to one or more correspondent elements in BMC. When there is a correspondence to two or three elements in BMC, any necessary decisions to choose on which specific element the sentence maps or on the separation in two or three statements, is responsibility of the analysts involved, whether based on the Business Model Ontology guidelines [6] or on their own business knowledge and heuristics.

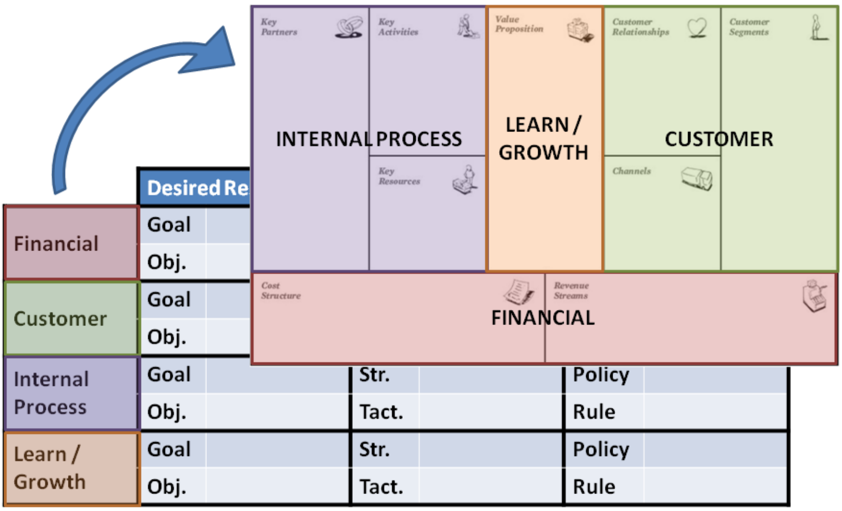


Fig. 3. BSC to BMC mapping in step 2

Surely there can be some overlaps between some elements, for example, the value proposition is closely related to any other element, especially the customer ones, so it is not impossible for mappings to occur outside elements other than the ones here proposed. Existing solutions in the business market propose slightly different mappings, but overall it depends on each particular case business type, and the informed or heuristic-based decisions of the analysts. All these matters should be the target for a final round of negotiations with the stakeholders, with the possibility for backtracking the changes made, back onto the UC.

Several iterations of this entire process should be performed until all parts are comfortable with the obtained solution. The generation of a BM through the use of this first PGR step serves two purposes: on one hand it allows to communicate with the stakeholders of the project in a more business-like language, in a format that is familiar to them; on the other hand, it allows for a direct alignment and enabled traceability between the UC elicited for the proposed IS and the BM to be analyzed and validated by the stakeholders.

4 Demonstration Case

The demonstration project is a new job matching and e-learning, cloud based platform, sponsored by technology-leading European companies, which aims to recognize and develop talents on the skills searched by employers, in order to tackle the shortage of professionals in technical areas. The core ambition is to offer targeted online education programs to improve ICT-skills, leveraging demand and supply on the European ICT job market, for science, technology, engineering and mathematics (STEM) people, preparing graduates for an industrial career and offering new skills and capabilities to empower current workforce.

The proposal is to create a Business Model that promotes creating a social network for e-learning-based talent matching, operating in Front-Office model, where each national office is responsible for its regional activities. UML Use Case diagrams, as a powerful and useful technique for capturing the system's requirements at the first high-level of abstraction, were used to elicit the functionalities and the entities that maintain and interact with the platform services. Regarding the project complexity and for the sake of our technique, just the five top-level Use Cases are considered to contribute to the generation of the high-level Business Model. After identifying the Use Cases (Fig. 4), the next step was to describe their behavior.

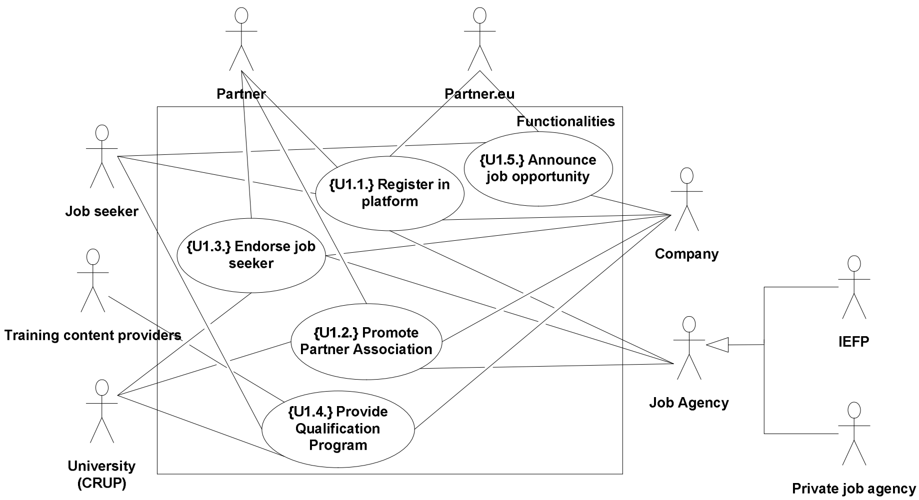


Fig. 4. Use Case Model regarding the Project Functionalities

Although diverse forms of information are available in the project (informal text, activity diagrams etc.), a structured description per Use Case during elicitation is not enough to generate the inputs for the Business Model as the stakeholders needed.

This due, it was decided to apply our developing technique in this live setting. As described in section 3, our proposal involves two steps which envision the filing of until twenty four (four BSC times six BMM) statements, not all mandatory, per Use Case, and a mapping to a BMC panel.

4.1 Step I – Inferring Goals and Rules from Use Cases

Although some projects require only the generation of the higher-level BM information, covering the filling of the three more abstract BMM items, in this project the analysts involved opted for performing the complete step, with all the six tasks. As an example we present all the abstract and concrete items filled with business information for the UC *{U.1.1} – Register in platform*, regarding the BSC Financial perspective (Fig. 5):

1. In UC *{U.1.1}* the goal to the Financial perspective includes the registration of the most possible amount job seekers on the platform;
2. The governing strategy is in the partnership between universities and employment agencies to reach a wider audience to register in platform;
3. Business policy indicates that all STEM ex-students with some specific requirements should be contacted for interest in registering in platform;

4.a) To achieve the associated Goal the objective is in register at least 1.000 job seekers annually;

4.b) To create the partnership to allow more registers the tactic is to be present and publicize at job fairs, universities meetings, etc.;

4.c) The associated business rule to the policy for contacting ex-students is that only one mail should be send to the ex-students/graduates to prevent spam practices.

		Desired Result		Course of Action		Directives	
{U1} Register in Platform	Financial	Goal	Register the most possible amount job seekers on the platform	Str.	Partner with universities and employment agencies to reach a wider audience	Pol.	All STEM ex-students and <6 months before graduation students should be contacted for interest in registering
		Obj.	Register at least new 1000 job seekers annualy	Tact.	Organize mailing lists to contact eventual job seekers Be present and publicize at job fairs, university meetings, etc.	Rule	Only one email should be sent to each student/graduate, so preventing spam practices Only STEM-origin, European-able workers should be registered.

Fig. 5. Financial Perspective with BMM items for *{U1} Register in Platform*

After populating the entire BSC grid for this first UC, in the remaining three perspectives, this first step iterated through the other four UC, populating all the BMM/BSC items accordingly. Following, the next step was to map these sentences to the BMC elements.

4.2 Step II – BSC to BMC Mapping

Although all six BMM items were elicited in the first step of this project, in this first round of the method execution, only the three more abstract items of each BSC perspective were considered relevant to be transposed to the BMC elements. This was due to the high-level positioning of the stakeholders to whom the BMC was addressed, and also to the shortage of information and ill-definitions of the project.

Again as an example, regarding UC *{U.1.1} – Register in platform*, its Financial perspective was categorized in the Cost Structure, the Customer in Channels, the Internal Process in Key Activities and the Learn/Growth in Value Propositions, all respectively. A discussion on sentence splitting to cover more specific aspects of dual elements, for example revenue-side information of cost-oriented UC, modification of the original sentences or their repositioning in different elements, is scheduled at a later stage within the stakeholders negotiation.

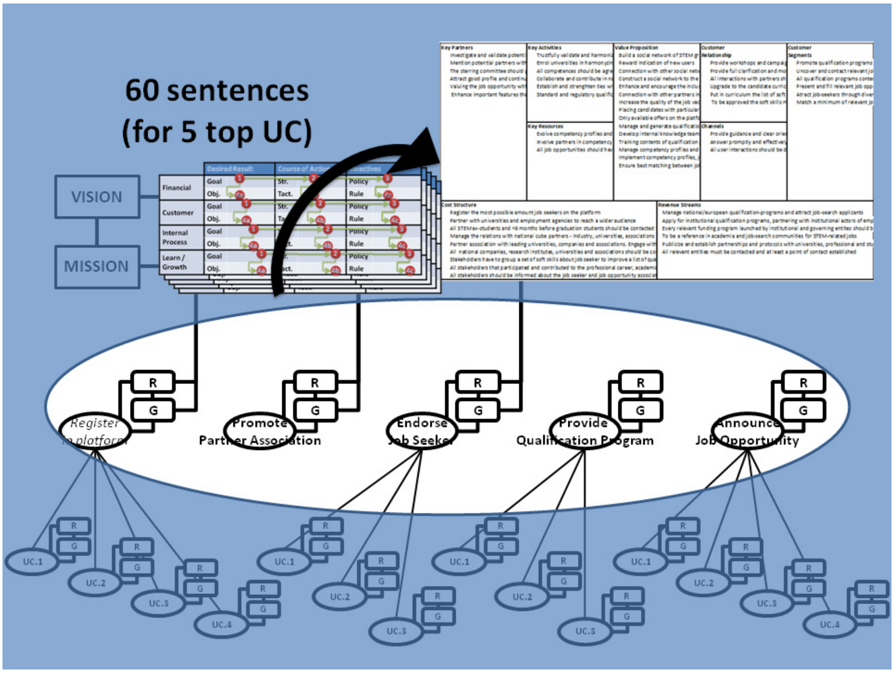


Fig. 6. Overview of the process for the current project

The solution BM for this project, after the first iteration, generated by aligning the resulting trios (PGR) with the BMC (Fig. 6), was presented to the project stakeholders. In the actual project phase, this information is currently being analyzed and also used as a token of discussion for the meetings adjourned.

5 Discussion

Our proposal for generating a Business Model through the elicitation of Business Goals and Rules from Process-level Use Cases has a dual standpoint, while it relies on 'traditional', established reference techniques and model representations; it also innovates on the organization and relationship of these to achieve a solid solution. The basis on established references strengthens the solution and the innovative organization allows for advances in research. Flexible and open solutions maintain an open door for dialogue between Business and IS/IT, leaving to analysts the choice for using other complementary techniques at some points, as also some room for negotiation with stakeholders.

The proposed method supports the effectiveness of BSC due to the individual connection to each Use Case and the associated elicitation of Goals and Rules elements, as they are segmented in the four BSC perspectives. More, it allowed for negotiating positioning of this previous information in BMC elements, and supported the Business-IS/IT dialogue by triggering interactions from the stakeholders. In this section we discuss the results obtained and the lessons learned from applying in practice our proposed method, and the perspectives for the future work ahead.

5.1 Results Obtained and Lessons Learned

One of the purposes for the proposal of this method was to overcome the lack of data in the initial documentation and more specific inputs from stakeholders to define the business requirements. The use of mixed techniques between Business and IS/IT allowed the analysts to, simultaneously, advance in the development of the IS and communicate with stakeholders, which helped in overcoming those issues. Its application represented a leap forward in the project, as it sparked the dialogue between the different stakeholders, which were in an entangled situation due to the lack of elements for communicating.

The players in the project are aware that creating a BMC alone will not translate into business success, but the solution obtained already had a positive impact in the development, sustainability and evolution of the project. The results have been promising so far, with positive feedback from involved stakeholders and research peers, but further work is needed in order to solve some issues and to validate the entire process.

On another point, according to the RUP referential, the types of tasks associated to these steps are to be performed by a person with a Business-Process Analyst profile, which seems too broad. In this project we observed that the first step requires a more IS/IT-oriented profile while the second step requires a more Business-oriented profile.

The amount of human resources needs to be taken into account, as there is the excess work of eliciting twenty four times more information for each UC (four BSC perspectives times six BMM items), which increases the complexity and length of the project. Nevertheless, if the resources are available, a comprehensive approach on the goals-rules tree [2] should be performed, ensuring a better validation of the root use cases, goals-rules associated information.

5.2 Future Work

Although properly guided, ahead we face issues around the openness of the associated elicitation techniques, as still much decision is left on the hand of the analysts, with the high-dependence on their Business-IS/IT knowledge and heuristics to be properly executed. Also, in dealing and connecting the technical elicited information with the abstract information of Goals and Rules, the mapping to BMC and the negotiation process, all need to be better explored and further analyzed.

The heavy reliance on the analysts' knowledge and their business heuristics calls for a stronger validation, being our intent to conduct further studies on the associated tasks in each step with diverse audiences, in real projects with Business/IS professionals and in educational context with students in the IS area. Also, involved in our own research group internal work, we plan to evolve the current steps to a complete PGR method, integrating with other internal methods and technologies, focusing on the Process-Goal-Rule trio associated information.

For now, the only tools used to aid in this process are some spreadsheets, but as the research evolves, development of an Eclipse-based tool is being considered. Linked to this, there is the need for greater formality on the process representation, especially for the lower-level items, where we will use SPEM. Also, there are issues regarding the use of natural language, opposite to formal approaches such as SBVR and its extensions [20]. The alignment and traceability issues will also profit with these added solutions, as the identification and formalization of the handled items makes it clearer and easier to connect and trace all business and technological elements.

6 Conclusion

Business Models are a top concern in todays Information Systems research, helping to link business and technological worlds, with Balanced Scorecard and Business Model Canvas as recurrent references in this topic. Also, the Process-Goal-Rule information trio is ever more interconnected and involved in issues of requirements elicitation, process modeling and business orientation and strategy.

Our work integrates all of these topics and proposes an innovative method to generate, align and maintain a Business Model for a desired Information System, based on the elicitation of Use Cases and its related Business Goals and Rules.

In this paper, we put forward a visualization to support the connection between a Business Model Canvas and the four perspectives of a Balanced Scorecard, eliciting BMM Goals and Rules for each problem space designed Use Case, building a strong focus on the business model strategy, and aiding in the future alignment between Business and IS/IT, and requirements traceability.

Acknowledgments. This work has been supported by *FCT – Fundação para a Ciência e Tecnologia* in the scope of the project: PEst-OE/EEI/UI0319/2014.

References

1. Zott, C., Amit, R., Massa, L.: The Business Model: Recent Developments and Future Research. J. Manage. 37, 1019–1042 (2011)
2. Salgado, C.E., Machado, R.J., Maciel, R.S.P.: Using Process-level Use Case Diagrams to Infer the Business Motivation Model with a RUP-based Approach. In: Information Systems Development, pp. 1–12 (2013)
3. Yu, E.S.K., Mylopoulos, J.: Using Goals, Rules, and Methods to Support Reasoning in Business Process Reengineering. In: Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, pp. 234–243 (1994)
4. Hickey, A., Davis, A.: A Unified Model of Requirements Elicitation. J. Manag. Inf. Syst. 20, 65–84 (2004)
5. Osterwalder, A.: The business model ontology: A proposition in a design science approach (2004)
6. Kaplan, R.S., Norton, D.P.: Linking the Balanced Scorecard to Strategy. Calif. Manage. Rev. 39, 53–79 (1996)
7. Kroll, P., Kruchten, P.: The rational unified process made easy: a practitioner's guide to the RUP. Addison-Wesley Professional (2003)
8. OMG: Business Motivation Model (2010)
9. Yu, E.S.K.: Towards modelling and reasoning support for early-phase requirements engineering. In: Proc. ISRE 1997 3rd IEEE Int. Symp. Requir. Eng., pp. 226–235 (1997)
10. Dardenne, A., Van Lamsweerde, A., Fickas, S.: Goal-directed acquisition. Sci. Comput. Program. 20, 3–50 (1993)
11. Burkhart, T., Krumeich, J., Werth, D., Loos, P.: Analyzing the Business Model Concept-A Comprehensive Classification of Literature. In: ICIS, pp. 1–19 (2011)
12. Supakkul, S., Chung, L.: A UML profile for goal-oriented and use case-driven representation of NFRs and FRs. In: Third ACIS Int. Conf. Softw. Eng. Res. Manag. Appl., pp. 112–119 (2005)
13. Yamamoto, S., Kaiya, H., Karl, C., Bleistein, S.: Goal oriented requirements engineering: trends and issues. IEICE Trans. Inf. Syst. 89, 2701–2711 (2006)
14. Amyot, D., Ghanavati, S., Horkoff, J., Mussbacher, G., Peyton, L., Yu, E.: Evaluating goal models within the goal-oriented requirement language. Int. J. Intell. Syst. 25, 841–877 (2010)
15. Horkoff, J., Yu, E.: Analyzing Goal Models – Different Approaches and How to Choose Among Them. In: ACM Symposium on Applied Computing, pp. 675–682 (2011)
16. Shao, J., Pound, C.: Extracting business rules from information systems. BT Technol. J. 17, 179–186 (1999)
17. Kardasis, P., Loucopoulos, P.: A roadmap for the elicitation of business rules in information systems projects. Bus. Process Manag. J. 11, 316–348 (2005)
18. Wan-Kadir, W.M., Loucopoulos, P.: Relating evolving business rules to software design. J. Syst. Archit. 50, 367–382 (2004)
19. Kapočius, K., Butleris, R.: Repository for business rules based IS requirements. Informatica 17, 503–518 (2006)
20. Milanovic, M., Gasevic, D., Rocha, L.: Modeling Flexible Business Processes with Business Rule Patterns. In: 15th International Enterprise Distributed Object Computing Conference, pp. 65–74. IEEE (2011)

How to Identify the Relevant Elements of "Context" in Context-Aware Information Systems?

Kurt Sandkuhl and Ulrike Borchardt

Institute of Computer Science, University of Rostock, 18051 Rostock, Germany
{kurt.sandkuhl,ulrike.borchardt}@uni-rostock.de

Abstract. Context-awareness is a feature of more and more applications, which adds further requirements to be taken into account in the implementation process. Though accepted approaches for software development exist, no accepted way for the inclusion of context has been established yet. An essential part of developing context based systems is to analyze and conceptualize the elements of the specific context required for the application under development, including their dependencies and mechanism of use. This activity of context modeling forms an important part of the system's specification, since it identifies relevant aspects of the application environment in a representation adequate for the modeling purpose. Within this paper we aim at closing this gap by introducing an approach for context modeling for the utilization in context-aware applications, providing a structure guiding through the process and illustrating it by examples as a reference for further projects.

1 Introduction

"Context-awareness" has emerged from a special and innovative feature of niche applications to a characteristic of many IT Systems in modern enterprises. At the beginning of the century, Dey's seminal work about context as information characterizing the situation of an entity [1] paved the way for context-aware ubiquitous computing and assistive systems. Nowadays, enterprise portals, groupware systems, assistive systems or control systems are including mechanisms to adapt to the users' situation on demand – to just name a few examples.

However, design and development of context-awareness in information systems still require substantial engineering work, i.e. there is no general development methodology for context-based systems. One reason for this probably is the variety of interpretations of the term context in the area of computer science (see section 2). An essential part of developing context based systems is to analyze and conceptualize the elements of the specific context required for the application under development, including their dependencies and mechanism of use. This activity of context modeling forms an important part of the system's specification, since it identifies relevant aspects of the application environment in a representation adequate for the modeling purpose. Furthermore, the context as such is also required during runtime of a context-aware information system, i.e. the context model is not only a conceptualization but has to be reflected in appropriate information structures and instantiated in the actual system.

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 290–305, 2014.

In this paper we focus on context modeling for IT-application cases in enterprises aiming at the support of human actors, i.e. we do not address context modeling for purely technical systems or cyber-physical systems. Based on experiences from several cases (see section 3) we propose to use techniques from enterprise modeling for context modeling and derive recommendations for context modeling activities. The main contributions of the paper are (1) an analysis of past context modeling cases with respect to commonalities in development processes, (2) an approach defining development steps of a context model, and (3) first experiences with the new context modeling approach.

As already indicated above, the question guiding our research is "What steps has a 'good practice' procedure for identifying and modeling the elements of context to include?" Good practice in this context has the meaning of a proven procedure for reliably completing a defined task, which originates from knowledge management [18]. The research process used in our work includes a deductive and an inductive phase. In the deductive phase, we analyze previous context modelling cases with respect to the sub-question "what commonalities and differences do context modeling procedures show?" The "data" available from these previous cases consist of project reports, notes of the developers and our own (undocumented) experiences. Based on these case data, an initial approach for a context modeling procedure was developed. The inductive phase includes usage and improvement of the initial approach in (new) application cases, which also serves validation purposes in order to reach the envisioned "good practice".

The remaining part of the paper is structured as follows: Section 2 provides a brief overview to context modeling in computer science and defines the term context. Section 3 presents context modeling projects from the past. These projects serve as experience basis for deriving our good practice approach for context modeling introduced in section 4. Section 4 describes the approach and relates it to the cases. Section 5 applies the new approach in an e-learning project. Section 6 summarizes the achievements and gives an outlook to future work.

2 Context Modeling in Computer Science

The term "context" has been used and still is subject of research in various application areas and sectors of computer science, which will be briefly summarized in this section. In the most general meaning, context describes what relates the entity under consideration to the environment surrounding this entity. What an "entity" is depends on the actual interpretation of context. Hoffmann [2] provided a way to classify these interpretations as follows:

- Linguistic context is used for disambiguating the meaning of words in texts and denotes the words surrounding the word of interest. An example is the approach presented in [3] to disambiguate keyword-based search using the paragraph surrounding a keyword of interest.
- Situational context includes any information characterizing the state or situation of a person, object or location. This information has to support the

purpose of understanding or being relevant for the interaction between user and application. Situational context models are often used in ubiquitous computing [1].

- Relational context includes any information pertinent to characterizing the relation of an entity to other entities, where this information is judged according to a given purpose. An example from problem solving is given in [4] using contextual graphs for this purpose.
- Formal representations of a perception or part of reality are like a model of an individual's viewpoint, which expresses a local view of the reality.

In this paper, we use the term context according to Dey, who defines context as "*any information that can be used to characterize the situation of an entity, where an entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and the application themselves.*" [1]

Although the term context is widely used in computer science, there is no general procedure how to develop context models. Many authors of context-based systems describe the way of developing the context model for their specific application, but do not provide a general view. [5] and [6] show examples for UML-context development in pervasive computing and OWL-based context for reasoning applications. Mena and colleagues [8] sketch a development process for context –aware systems and identify invariant characteristics of context as part of their work. These characteristics are (a) context relates always to some entity, (b) is used to solve a problem (c) depends on the domain and (d) is a dynamic process.

The development of a procedure for context modelling should be anchored in experiences from past projects in order to exploit the lessons learned. In addition to our own experiences presented in section 3, surveys analysing previous context work form an important contribution. Bazire and Brézillon [7] analyzed the state of research in the area including more than 150 definitions. One of the conclusions from this work study is that "*the context acts like a set of constraints that influence the behaviour of a system (a user or a computer) embedded in a given task.*" Furthermore, in [7] it is concluded that context definitions should be analysed using six parameters: constraint, influence, behaviour, nature, structure and system.

3 Context Modeling Cases

The context modeling approach presented in section 4 is based on experiences from a number of projects where adaptivity to situations was a decisive feature. Two of these context modeling application cases are presented in this section with their background or application scenario, the development process for the context model, and the context model as such. The selected context modeling cases are from information demand context and context-based ontology matching. The other cases analyzed but not included for brevity reasons are context-modeling for capability-as-a-service [14], situation-based messaging [15] and decision support [16].

3.1 Information Demand Context

Background. The first context modelling case originates from the field of information logistics which focuses on improving information flow in enterprises and on demand-oriented information supply. A core subject of demand-oriented information supply is to capture the needs and preferences of a user in order to get a fairly complete picture of the demand in question. This requires an understanding of what information demand is and a method for capturing and analysing information demand.

Information demand has a strong relation to the context in which such a demand exists [8]. The organisational role having the demand, the task the information is demanded for as well as the setting in which such tasks are performed are important aspects for understanding information demand. Thus, the concept of information demand context has been defined as follows: "*An Information Demand Context is the formalised representation of information about the setting in which information demands exist and comprises the organisational role of the party having the demand, work tasks related, and any resources and informal information exchange channels available, to that role.*"

Development Process of Context Model. The approach used for developing an information demand context model is information demand analysis; the main characteristics are published in [4] and summarized in the following.

Scoping is the process of defining the area of analysis and is done with the purpose of selecting parts of an organisation to be subjected to analysis. This phase also includes the identification of the roles (individuals) relevant for the continued information demand analysis. *Information Demand Context Modelling* is mainly performed through participative activities such as joint modelling seminars where the participants themselves are involved in the actual manufacturing of different models. This process is usually supported and facilitated by a method expert who could be an internal or external person. The key to context modelling is to identify the interrelationship between roles, tasks, resources and information. No regard is given to the sequence of activities, resource availability, etc. *Information Demand Context Analysis and Evaluation:* Once the necessary knowledge about the information demand contexts is obtained, it can be used for a number of different purposes. One purpose is evaluation where different aspects of information demand can be evaluated in relation to roles, tasks, resources and information. It is also suitable to address the results from the modelling session with respect to motivation and purposes expressed during scoping activities. Focusing on information demand contexts provides only an initial view of information demand without any consideration given to such aspects as individual competence, organisational expectations and requirements in terms of goals, processes etc. Depending on the intentions behind the analysis further activities might be required. The method provides a number of method components supporting such activities. If the method user wishes to investigate such additional aspects of information demand, he or she can do this by using subsets of the other methods, notations and languages.

Context Model. An information demand context model basically is an excerpt from an enterprise model for a specific role showing the processes the role is involved in, the co-workers in the organization structure, the resources used and the IT-systems applied.

Figure 1 illustrates the relation between enterprise model and information demand context. For the considered role, the information demand context is derived from an existing enterprise model by extracting (a) all tasks the role is supposed to perform or is responsible for, (b) the co-workers, superiors and sub-ordinate positions, (c) the resources required for the tasks gained in step a. The information demand context at design time is instantiated with the active tasks, persons assigned to the positions and resources in use and forms the actual information demand at the point of runtime.

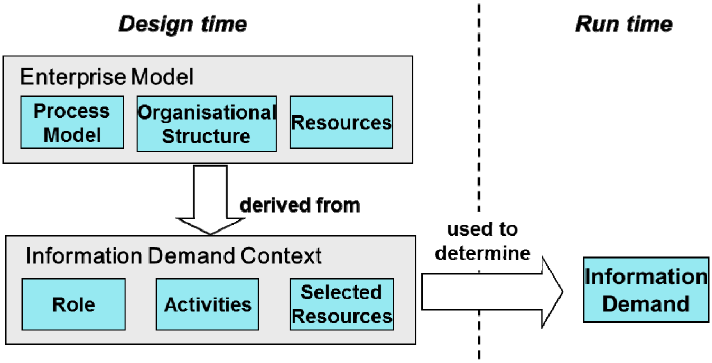


Fig. 1. Information Demand Context

3.2 Context-Based Ontology Matching

Background. Although ontologies are developed for various purposes and domains, they often contain overlapping information. Ontology matching aims at finding similar entities or translation rules between two ontologies. Ontology matching is an important technique for creating a collaborative semantic web. However, currently existing approaches for automatic ontology matching do not sufficiently take into account context dependencies in the process of matching. This leads to situations where the results of automatic matching are of limited or no use for the task or application at hand. An increased user involvement can be a way to improve the quality of matching results [9]. The second context modeling case aims at facilitating a new way for user involvement by using a context ontology capturing both, tasks of the user and user preferences.

Development Process of Context Model. The context aims at reflecting the information demand of a role in an enterprise. Role here means a part of a larger organizational structure clearly defined by the responsibility it has within that structure [10]. The context is modeled in two levels: abstract context and operational context. Abstract context is an ontology-based model integrating information about the role. Operational context is the instance of the abstract context for a specific role. Normally the context consists of three parts:

- The information about the tasks of a role included in the enterprise ontology.
- Information about tasks of the role that is related to the enterprise but not inside the enterprise ontology. This is additional information provided by the role based on his/her knowledge.
- Additional information about the role, every individual having the role, for example, the competence of the individual having the role.

Context Model. An example for a context model for context-based ontology matching is shown in Figure 2. The example shows the context model for a person (concept “expert=person(CODISPLAY)”) with the projects and training courses this person was involved in. This context model is supposed to represent the competence of the person which should be taken into account when searching for experts and competences complementary to the person’s competence. The context model is used in ontology matching during relevance calculation of matching concepts in two ontologies identified for expert finding and competence supply purpose.

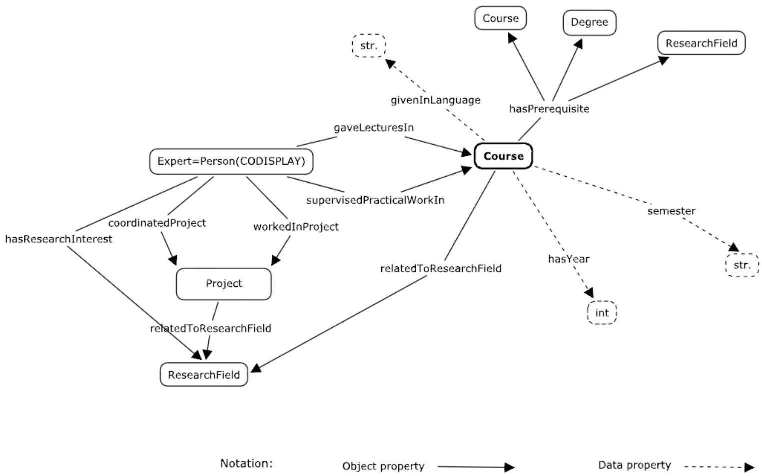


Fig. 2. Ontology Matching Context

4 Context Modeling Approach

Based on the cases presented in section 3, an approach for context modeling has been devised. The cases were analyzed in order to find commonalities and differences in the development process and in structure and content of the context models. The most significant commonalities were (a) all cases required to understand what user activities were supposed to be supported using context, (b) the context models were reflecting the potential variations to be supported and (c) the common phases of engineering projects (e.g. requirements elicitation, design, implementation, test) were represented in all cases. The most significant differences were the context models, i.e. their structure and content showed hardly any commonalities. Our conclusion is that we have to support both, developing context models for applications (i.e. context model types) and context models for specific usage cases of these applications (i.e. actual context models). We propose a context modeling approach consisting of six steps, which are discussed in the following:

1. Scenario modeling for the future context-based application
2. Variability elicitation
3. Initial context model development
4. Implementation of context-based application
5. Alternating model-instance improvement
6. Theory and practice validation

These six steps form our hypothesis for a context modeling approach and require evaluation regarding completeness, practicability and refinement needs. They reflect the typical phases of engineering projects found in all analyzed cases with a specialization for accommodating variability aspects and different context types.

4.1 Scenario Modelling

The purpose of the first step is to identify user groups and intended scenarios of use for the future context based system. This step is similar to the first phase of information system development or software projects. In order to understand which user groups exist and how their ways of using the future system differ from each other, the process supported by the system, the information input and output, possible connections to other systems and processes, or the integration of resources have to be analyzed and described. This may be done using conventional use case modeling (e.g. from RUP), business modeling in UML 2.0, goal-process-actor modeling in 4EM or other techniques. The result of this step are scenario representations, e.g. as diagrams or visual models. In order to be suitable for context modeling, the scenario descriptions have to include and identify:

- The different user groups of the future context-based IT system
- The tasks the users are supposed to perform with the future system. This should at least include the primary scenario (often referred to as “success scenario” or “happy flow”) with steps to perform
- Information input or conditions which cause branching in the flow of actions during the tasks

4.2 Variability Elicitation

The second step is probably the most important one. A context model has to include in what situations and on what inputs or events what kind of adaptations in the context-based system should be made. Adaptations can concern functionality, behavior, output or appearance of the system. Since the results of these adaptations of the system can be considered as variations of the use of the system, the system's behavior or even the system's configuration, it is decisive to understand the cause and kind of the variation. In order to determine cause and kind of variation, two aspects have to be investigated: the variation aspects and the variation points.

Variation Aspects. Variations in behavior, functionality or content of context-aware systems can be caused by different aspects, like the user groups, the task performed, the information input, etc. In order to identify the relevant aspects, the scenario models developed in step 1 have to be analyzed. In principle, different strategies of doing so are possible, like investigating all tasks in the scenario models and their variations or focusing on causes for different branches in the scenario. Since deciding on the best strategy would require more cases and data collected from analyzing them, we propose a "brute force" strategy based on the scenario models:

The modeling languages mentioned in step 1 "Scenario modeling" include different model component types, like the "process" or "actor" types in 4EM [11]. For each of these model component types, it has to be examined whether different instances of this component type would require an adaptation in the context-aware system. For those component types causing an adaptation it has to be investigated what characteristic of the component type actually is decisive for the adaptation. If, for example, "process" component types would cause adaptation, it has to be investigated whether this is due to process input, process output, process duration or other characteristics. The identified component types and their decisive characteristics are called variation aspects.

Variation Points. Within each variation aspects, the variation points define under which conditions or for which events an adaptation in the context-aware system has to happen. Often even the kind of adaptation can be identified together with the variation points. In order to identify the variation point, all variation aspects identified in the above procedure have to be examined. It is recommended that this is done based on the scenario models by assuming alternatives in the scenarios regarding the validation aspects under consideration.

4.3 Develop Initial Context Model

According to the definition of context, the context contains all information characterizing the situation of an entity. We assume that this information consists of different elements and that each element has different attributes. An example would be a context element "user group" with the attribute of "list of user groups to be distinguished" and "individuals assigned to the user groups". For developing the initial context model, the first task would be to define a context element for each of the identified variation aspects and to decide on the attributes for the context model.

The second task aims at investigating what type of adaptation of the context-based application is related to each context element. For this purpose, we assume that a context-aware application not only has to adapt its own behavior with respect to functionality or what information is provided (active role) but also needs to provide information to other "context-aware" components outside the context-based application to be developed (passive role). An example would be a context element "current user location" which can be used to adapt the context-based application under development, but which also serves as input for other applications using location information.

When investigating the type of adaptation related to context elements, this passive vs. active role of the context and the content vs. application orientation of the context can be used as aid. For all context elements identified, the following questions should be answered using the variation points from step 4.2:

- Does the context element influence the behavior of the context-based application only or also an external "context-aware" applications? (active role; internal and/or external)
- Does the context element influence the information provision or the application behavior or both?
- Do the attributes of the context element have to be updated by the context-based application only or could there be a need to also use external "context-aware" applications?

The above questions would result in a classification for each context element on the one hand side into internally relevant and updated in the context-base application or (also) externally relevant and updated. On the other side there is a classification into relevant for behavior adaptation or relevant for information provision. This classification helps during software design of the context-aware application for deciding on operations on context elements and their external visibility and related interfaces.

4.4 Implementation of Context-Based Application

The next step in our approach is the implementation of a context-based application using the initial context model from 4.3. This step basically is not elaborated in our approach, since it usually includes a software development process and many software development approaches exist which could be integrated (see [17] for an overview). From the software development process viewpoint, context modeling can be considered as part of the requirements specification task or as part of the early software design task.

However, using the context model for implementing an application based on it is part of the validation of this model and will give valuable and necessary feedback regarding required improvements and utility of the model. This is why the "link" to software development was included as an explicit step in our approach. As a result of this step, experiences from using the context model including improvement requirements or a confirmation of the context model's utility are expected.

4.5 Iterative Improvement

As discussed in section 2, context models are used in various application domains with different needs, i.e. different context types exist, some of these types are used in several applications and we expect more types and applications for these types to emerge in the future. Thus, development processes for context models include the development of the type of a context model – in case this does not yet exist – and development of actual context models of this type. In order to reach a high "fitness for purpose" of the context model types and utility of the actual context models we propose an iterative approach, which resembles Boehm's well-known "spiral model" [12] for development of software systems and was inspired by this work. Advantages attributed to the spiral model are early validation and continuous improvement of artefacts developed in the process. From our perspective, these characteristics are very useful for collecting feedback from developers of context-based applications and reaching a high maturity of context models. Our proposal is to develop context models in an iterative way consisting of alternating development and validation steps for both, model type and actual model. Figure 3 illustrates the overall approach.

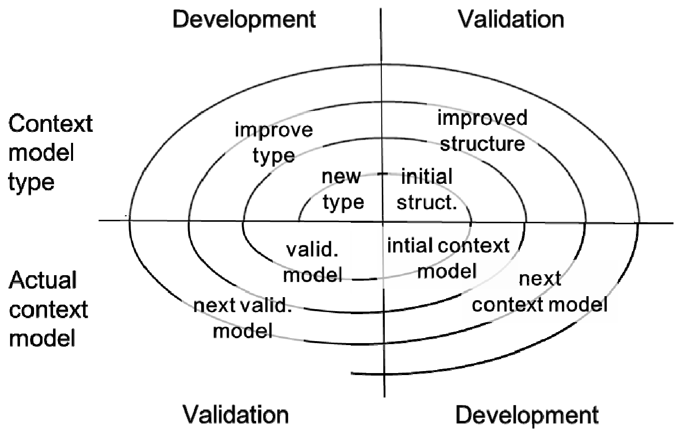


Fig. 3. Iterative process of development and validation

In this context, the initial development of a context model type (step 1; marked as "new type" in Figure 3) is performed according to the steps described in sections 4.1 to 4.3. The next step should be the validation of the initial structure (step 2; "initial struct." in Figure 3) consisting of checking internal consistency and soundness. The next step recommended is to apply the context model type by building an initial version of the context-based application, i.e. the next step includes developing an actual context model (step 3, "initial context model"). Afterwards, the initial model also has to be validated by using the context-based applications (step 4, "valid. model"). Steps 1 to 4 form the first iteration. The second iteration would then improve both, type and actual model.

4.6 Theory and Practice Validation

The validation of context model type and actual model has to be performed including developers and users of context-based applications, and encompassing both theory and practice. Among the many scientific approaches for validating qualitative research results, we base our proposal for validation activities to be performed on the work of Lincoln and Guba [13, p. 289 ff.] on "naturalistic inquiry". On the one hand, we distinguish between theoretical and practical validation. Theoretical validation means assessing an approach within the theories of the domain the approach is part of or supposed to contribute to. For context type validation, this means to assess the soundness, feasibility, consistency within the body of knowledge in, for instance, computer science and information systems. Practical validation encompasses all kinds of application of the context model for validation purposes, which requires defined procedures and documenting results. This could be simple lab examples illustrating the approach, controlled experiments in a lab setting, application in industrial cases, etc.

On the other hand, we consider the context of validation and distinguish between validation by the developers of the approach in their internal environment, validation by the developers outside the internal environment, and validation by other actors than the developers. Combining these two perspectives leads to a two by three matrix, which is depicted in Table 1. The cells of this table show typical ways of validation for the different combinations of the two perspectives.

Table 1. Proposed validation steps for context models

	Theory	**Practice**
Internal, development team	Validation against state of research	Prototype implementation, test in lab environment
External, in validation context	Peer-review, comparison to known best practices	Case studies for evaluation purposes
External, in application context	Development of extensions by external actors	Use of the artifacts developed for solutions

Using the above matrix, the different iterations of the context model development described in section 4.5 should proceed from theory to practice and internal to external validation. Thoroughly validated context types will include all parts of the matrix and involve several iterations.

5 Application in KOSMOS Project

The context modeling approach presented in section 4 has been applied in the KOSMOS project in order to validate feasibility and usefulness, and to gather first experiences and hints how to improve it. The KOSMOS project aims at attracting new target groups to university education and to develop and explore new study formats. New target groups and formats need an adjusted or different kind of support by learning management systems compared to the traditional target groups, since didactic and pedagogical concepts also differ. In order to facilitate this adaptivity requirement, our approach is that learning management systems (LMS) should be flexibly adaptable to the learner's individual demands when it comes to contents and applications supporting the learning process. We consider a context-aware LMS a suitable technical implementation of this requirement. In KOSMOS, this LMS is supposed to be a portal integrating existing and future learning objects and tools supporting the different learning phases. This portal is called "myKOSMOS". The development process of myKOSMOS is performed according to the steps introduced in section 4 and is described in the following. However, the application of myKOSMOS is not subject of this paper.

Step 1: Scenario Modeling

For this first step of scenario modeling we used an approach from enterprise modeling based on Troux Architect as a tool and Troux Semantics as notation. Consequently, we modeled the different planned ways how myKOSMOS would be used by the future users. This resulted in process model-like scenarios, as depicted in Figure 4 showing the example "assignment work in distance learning".

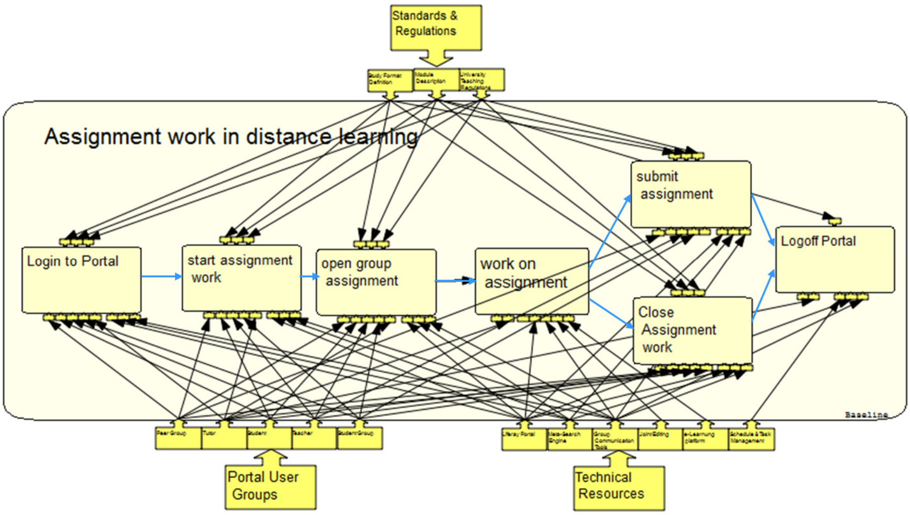


Fig. 4. Scenario model for "assignment work in distance learning" to be supported by the context-based portal myKOSMOS

Step 2: Variability Elicitation

The second step is to identify variation aspects and variation points. As a starting point for this activity, we briefly describe the scenario in order to show the variations: the scenario starts with the student logging in. According to his profile he is provided with an individually configured entry page, making offers for his learning process. The variation here is due to his study format as well his individual preferences captured during different sessions. Following his course of study, completing different modules within the study format, the student choses to open or proceed with his assignment work for a certain module, which is loaded presenting the recent state of his work in progress. Once having caught up with his recent results, the student is confronted with different tasks to be fulfilled in order to fulfill the assignment, however is free to choose which task to pick. A regular assignment the designed study formats includes information research the portal supports providing the appropriate sources for the study format. In addition many assignments also involve the communication with fellow students since they are assigned group work. In the process the work should be documented to be handed in, where the kind of documentation being determined in the assignment description. During the work process coordination issues between the team members should be resolved as well, which might be due to the individual time tables and working hours, as well as the specific interests or responsibilities within the task assignment. At the end of each session the user has the choice between submitting his work for the correction process and simply closing it to proceed in another sessions. The consequence at the end of the session would be a log-off which is accompanied by a profile update due to the user's behavior during the session.

Variation aspects. According to the above description, the following component types caused variations: activity (for capturing the portal use processes), user group (like assignment group, study format group), documentation type (Word document, interactive document, learning journal entry etc.), application type (communication support, groupwork support, editing support, search, etc.). For all component types, the characteristics of the type decisive for the variation is the actual instantiation, i.e. what user group logged in, what type of documentation is used etc.

Variation Points. The scenario model for myKOSMOS included the following variation points: login to portal (variation due to user group and study format of the user logging in), start assignment work (variation due to the type of assignment given in the module description), open assignment (variation due to the actual student group working on a specific assignment), work on assignment (variation due to status of the actual work) and open group assignment (variation due to the tool support available).

During analysis of the entities we discovered that also the need to distinguish the type of learning task (e.g. assignment, exercise, lecture, etc.). These tasks were represented by the different scenario models.

Step 3: Develop initial context model

After identifying the variation aspects and variation points we combined them into the necessary context elements and their classification within the scenario as described in section 4.3. The result of this step is provided in Table 2.

Table 2. Context Elements for myKOSMOS

Context element	*Context element attributes*	*Type of adaptation*	*Originated from variation aspect*	*Originated from variation point*
Study format	Name, modules assigned	Active, content and behavior	Type: portal user group, char.: study format	“Login to Portal”
Module	name, assignments	Active, content and behavior	Type: standards & regulations, char.: module description	“Start assignment work”
User Group	Name, members, assignments, Preferences	Active/passive, behavior and content	Type: Portal user group, char.: student group	“Open group assignment”
Assign-ment	Description, deadline, type	Active/passive, content and behavior	Type: learning tasks, char.: assignment	“start assignment”
Appli-cation support	Description	Active/passive, content and behavior	Type: activities, char.: kind of activity	“work on assignment”
Prefe-rences	Application type	Active/passive, behavior	Type: technical resources, char.: group commu-nication tools	“Open group assignment”

Step 4: Implementation of context-based application
Using the initial context model with its identified elements, we started implementation phase of the portal myKOSMOS. The main effort of this phase lies in transferring the context model into a data model. The implementation takes places via a Liferay development which will be extended by a context processing component.

Step 5: Alternating model-instance improvement
Following the initial implementation of myKOSMOS a validation is necessary. Already in the process implementation minor adjustments are done due to implementation specifics, as e.g. the preferences as such have to be refined to be captured from the behavior of the user. This part of the improvement means successive refinements by the concretization of the scenarios under the implementation process.

Step 6: Validation phase
Finally the rigorous validation of the context model is necessary. Referring to the validation phases as mentioned in section 4.5 the internal and theory related validation is finished with the end of the modelling. The internal and practice related validation is ongoing being closely connected with the successful implementation of the finished context-aware application.

6 Conclusion

The goal of the paper was to develop a 'good practice' procedure for identifying and capturing all necessary context elements for context-based applications. The approach presented in section 4 was derived from an analysis of previous context modeling cases and has gone through an initial validation in the myKOSMOS project. The approach is based on scenarios and can be run in different manners. One example is to start with one initial scenario, and build a context following this one scenario only. Another example would be to add further scenarios, which makes the context more complex from the beginning, but certainly allows for a more extensive validation.

Our experiences during the validation of the method by application in the KOSMOS project included that the differentiation between variation aspects and points turned out to be most difficult for the creation of the context model. Furthermore the transfer of the context model to an explicit data structure suitable for implementation was labor-intensive using the style of modeling with Troux Semantics as shown here.

Future work will have to include theoretical and practical aspects. From a theoretical perspective, we aim at further formalizing the concept of variation points and variation aspects and how the variation has to be reflected in the context model. The classification in content and behavior aspects and in internal and external effects seems useful, but is not yet clear enough. Furthermore, the transition from the context model into an implementation of the model has to be further investigated with the objective to support the design of software components implementing the context concept and the envisioned behavior captured in the context model. Furthermore, the integration of the context modeling approach and software engineering processes should be further investigated. Since a model representation of context can be part of the early design of the information system to be developed, it might also influence the architecture of the overall system. Furthermore, the model carries requirements which need to be taken into account during the development process. These aspects need further exploration.

From the practical perspective, implementation of the myKOSMOS context component, using it in selected study formats and collecting improvement potential and experiences during the usage will be an important future activity. In accordance to steps 4 and 5 of our context modeling approach, we aim for finishing the first complete iteration of type development and implementation before starting an improvement cycle and we expect to collect sufficient experiences from internal practical validation to be able to continue with external validation.

Acknowledgements. The work presented in this paper was supported within the project KOSMOS (Konstruktion und Organisation eines Studiums in Offenen Systemen) funded by the BMBF (Federal Ministry of Education and Research, Germany) and the European Social Funds of the European Union.

References

[1] Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing 5(1), 4–7 (2001)

[2] Hoffmann, P.: Context-based Semantic Similarity across Ontologies. PhD-thesis. Claude Bernard University of Lyon (December 2008)

[3] Finkelstein, L., Gabrilovich, E., Matias, Y., Rivlin, E., Solan, Z., Wolfman, G.: Placing search in context: the concept revisited. In: Proceedings of the 10th International Conference on World Wide Web, pp. 406–414 (2001)

[4] Brézillon, P.: Context Modeling: Task Model and Practice Model. In: Kokinov, B., Richardson, D.C., Roth-Berghofer, T.R., Vieu, L. (eds.) CONTEXT 2007. LNCS (LNAI), vol. 4635, pp. 122–135. Springer, Heidelberg (2007)

[5] Henricksen, K., Indulska, J., Rakotonirainy, A.: Modeling context information in pervasive computing systems. In: Mattern, F., Naghshineh, M. (eds.) PERVASIVE 2002. LNCS, vol. 2414, pp. 167–180. Springer, Heidelberg (2002)

[6] Wang, X., Zhang, D., Gu, T., Pung, H.: Ontology based contextmodeling and reasoning using owl, pp. 18–22 (2004)

[7] Bazire, M., Brézillon, P.: Understanding Context Before Using It. In: Dey, A., Kokinov, B., Leake, D., Turner, R. (eds.) CONTEXT 2005. LNCS (LNAI), vol. 3554, pp. 29–40. Springer, Heidelberg (2005)

[8] Ben Mena, T., Bellamine-Ben Saoud, N., Ben Ahmed, M., Pavard, B.: Towards a Methodology for Context Sensitive Systems Development. In: Kokinov, B., Richardson, D.C., Roth-Berghofer, T.R., Vieu, L. (eds.) CONTEXT 2007. LNCS (LNAI), vol. 4635, pp. 56–68. Springer, Heidelberg (2007)

[9] Shvaiko, P., Euzenat, J.: Ten challenges for ontology matching. In: Meersman, R., Tari, Z. (eds.) OTM 2008, Part II. LNCS, vol. 5332, pp. 1164–1182. Springer, Heidelberg (2008)

[10] Levashova, T., Lundqvist, M., Sandkuhl, K., Smirnov, A.: Context-based modelling of information demand: Approaches from information logistics and decision support. In: Proceedings of the 14th European Conference on Information Systems (2006)

[11] Sandkuhl, K., Stirna, J., Persson, A., Wissotzki, M.: Enterprise Modeling– Tackling Business challenges with the 4 EM method. Springer, Heidelberg (2014)

[12] Boehm, B.W.: A spiral model of software development and enhancement. IEEE Computer 21(5), 61–72 (1988)

[13] Lincoln, Y., Guba, E.: Naturalistic Inquiry. Sage Publications (1985)

[14] Sandkuhl, K., Koc, H., Stirna, J.: Context-aware Business Services: Technological Support for Business and IT-Alignment. In: 5th International Workshop on Business and IT-Alignment @ BIS 2014. Accepted for publication in LNBIP Workshop proceedings (To be published in October 2014)

[15] Meissen, U., Pfennigschmidt, S., Sandkuhl, K., Wahnfried, T.: Situation-based Message Rating in Information Logistics and its Applicability in Collaboration Scenarios. In: Euromicro 2004 Special Session on "Advances in Web Computing", August 31-September 3. IEEE Computer Society Press (2004)

[16] Sandkuhl, K., Seigerroth, U., Tarasov, V.: Context Modelling for Decision Making in Competence Supply: Experiences from a Military Case. In: Villa-Vargas, L., Sheremetov, L., Haasis, H.-D. (eds.) Operations Research and Data Mining ORADM 2012 Workshop Proceedings, Cancun, Mexico. Mexican Petroleum Institute, Mexico City (2012) 978-607-414-284-6

[17] Sommerville, I.: Software Engineering, 9th edn. Pearson (2010) ISBN: 9780137053469

[18] Davenport, T.H., Prusak, L.: Working Knowledge. Harvard Business School Press, Boston (1998)

Identifying Best Practices in Business Process Management Using Fuzzy Analytical Hierarchy Process

Nazli Shahim and Charles Møller

Center for Industrial Production,
Department of Business and Management,
Aalborg University

Abstract. To implement Business Process Management several best practices can be distinguished. This paper provides an overview of the best practices that can support BPM practitioners to develop business process maturity. Various common BPM practices are derived from a literature survey and interviews with process architects. Practices are assessed through questionnaires and final best practices selected through Fuzzy Analytical Hierarchical Process are presented. The study has been carried out among process architects in 3 large telecommunication companies in Denmark. Although every organization has different needs, the best practices identified in this paper are believed to have a wide applicability across various telecommunication organizations. The overall conclusion is that the framework is indeed helpful in supporting BPM application improvement and can be put in practice by the BPM practitioner community.

Keywords: Business Process Management, Best Practices, Fuzzy AHP.

1 Introduction

Globalization, changed technologies, new rules and the erosion of business limitations are always the motivating factors that drive changes and improvement approaches. Organizations need to respond to global competition, demanding customers and employees and the like by becoming faster, more flexible and more focused on competition, customers, quality, time and processes. Various studies have found that to address the competitive environment issues, less emphasis must be put on the functional and hierarchical organizations. The attention must be paid to the whole chains of processes and to mature them instead [1]. According to many researchers, process orientation has brought gains and increased efficiency and competitiveness within the organizations [2-4]. In virtually every industry, companies of all sizes have achieved extraordinary improvements in cost, quality, speed, profitability, and other key areas by focusing on, measuring and redesigning their customer-facing and internal processes. Process orientation has been known for a long time now. Hammer and Champy, as the experts in Business Process Reengineering (BPR) define business process as actions that result in customer satisfaction by taking one or several inputs [5]. Introduction of process orientation brought with it the need for new disciplines

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 306–319, 2014.

like BPR and Business Process Management (BPM). BPM is about tools, methods and techniques with the aim of supporting the design, organizing, management, and analysis of operational business processes. Positive results of BPM are more often reported than the negative ones. The table below shows some of the results of BPM according to various authors.

Table 1. BPM reported results

Author	Results
[6]	Customer satisfaction
[7]	Throughput satisfaction improvement, declined process costs, progresses in the quality of business processes, improvement in delivery reliability
[8]	Positive effects on standardization, having a common language, customer satisfaction and cycle time
[9]	Better financial performance
[10]	Focus on customers regarding the linkage among all the main activities, guaranteeing continuous discipline, stability of quality performance
[11]	Reduced internal functional conflict and improved business performance
[4]	Improved their financial and nonfinancial performance
[12]	Lead time reduction, decreasing costs, and more progress in the consistency of the results
[13]	Increase in cost

As this research focuses on the best practices in BPM application, the first step was to identify the most common BPM elements. A broad list could be provided for BPM elements but the most common core elements seen in BPM implementing are information technology, culture, strategic alignment, methods, people and governance [14-16]. The next step was to find the common practices related to these core elements which were obtained through literature survey and interviews with process architects in three telecommunication companies in Denmark. The last step was to identify the best practices. To this end, we have again contacted process architects to fill out the questionnaire. The data gathered through questionnaires were assessed via Fuzzy Analytical Hierarchy Process (FAHP). Application of the FAHP has been very useful in different study fields in prioritizing among several choices by decision-makers to arrive at a consensus decision [17]. The results of this study are believed to serve as guidance to which elements and practices should be focused on when implementing and improving BPM.

1.1 Aims of Study

This paper seeks to provide BPM practitioners and academics with insight into the most popular heuristics to derive improved PBM implication.

1.2 Research Structure

The paper consists of several parts. The theoretical part is about BPM, AHP and fuzzy AHP (FAHP). The methodology part comprises data collection method and the steps of the FAHP for the analysis (in the appendix link given later in this paper) and finally indication of the results.

2 Theoretical Framework

The theoretical framework deals with explanations and definitions of the several elements that this research works on. This part gives the theoretical understanding of BPM, AHP and FAHP.

2.1 Business Process Management

DeToro and McCabe described BPM, a different way of managing an organization in comparison to hierarchical one [18]. Pritchard and Armistead also reported similar description as they viewed BPM as a holistic approach for managing the organizations [19]. BPM aims at bringing in benefits like efficiency, agility throughout the organization whilst making sure that the goals of the organization are met. Zairi declared that BPM is about change, implementation of better systems in the organizations and an approach to reap competitive advantage within the market [10]. The Zairi's view is supported by Spanyi who also confirmed the changed focus of BPM and stated BPM is actually about cultural change [20]. This paper focuses on BPM core elements and the related common practices and contributes to application of FAHP on identification of BPM best practices within the selected telecommunication companies. The common practices for each element listed below are gathered trough literature survey [21-25] and interviews with process architects:

2.1.1 Information Technology

Information Technology (IT) refers to the software, hardware, and information systems that enable and support process activities [22]. The related common practices are as follows:

- There are BPM software tools available supporting the process modeling, simulation, analysis and reporting.
- There is "process aware" IT enabled implementation tool which automate the transformation of process models to executable processes.
- There are software tools available which assist in automating visualization, control and management of the running processes.
- Automated process improvement tools are available which can alter the business processes, for example, self-tuning/self-learning tools.
- There are BPM project management software tools available to facilitate the BPM initiative.

2.1.2 Culture

Culture, refers to the collective values and beliefs that shape process-related attitudes and behavior to improve business performance [26]. The related common practices are as follows:

- The organization manages and analyses its ability to accept process change and adaptation, even when process cross departmental boundaries.
- There is a common belief within the organization of the value of BPM and its ability to benefit the business, and is that belief promoted well.
- The attitudes and behaviors of those involved in the BPM initiative are towards process improvement throughout the whole organization - are these attitudes and behaviors managed well.
- There is a managed active commitment from senior executives towards the BPM initiative.
- The encouragement of "BPM Community" social networks within the organization is managed.

2.1.3 People

While the information technology factor covered IT-related resources, the factor “people” comprises human resources. This factor is defined as the individuals and groups who continually enhance and apply their process and process management skills and knowledge to improve business performance [27]. The related common practices are as follows:

- There are BPM process skills and expertise associated with the individual’s role and responsibility.
- Process management knowledge is towards process management methods and best practice understood along with the impact these have on the organization.
- There is on-going development and maintenance of education and learning towards process skills and knowledge.
- There is encouragement of, and analysis of, process collaboration between individuals within the BPM initiative.
- Willingness of individuals is within the BPM initiative to lead, to take responsibility and accountability for business processes encouraged and managed within the organization.

2.1.4 Methods

Methods, in the context of BPM, have been defined as the tools and techniques that support and enable consistent activities on all levels of BPM (portfolio, program, project, and operations). The methods dimension focuses on the specific needs of each process lifecycle, and considers elements such as the integration of process lifecycle methods with each other and with other management methods, the support for methods provided by information technology, and the sophistication, suitability, accessibility, and actual usage of methods within each stage [21]. The related common practices are as follows:

- Documentation of methods for identifying and modeling "as is" and "to be" processes.
- Existence of methods in place relating to process implementation and execution.
- Defined and documented methods available relating to the collection of process metrics to be used in the measurement and control of the process designs.
- Existence of method for process improvement.
- Methods to assess (and improve) the project management and overall approach of the BPM initiative.

2.1.5 Governance

BPM governance is dedicated to appropriate and transparent accountability in terms of roles and responsibilities for different levels of BPM (portfolio, program, project, and operations). Furthermore, it is tasked with the design of decision-making and reward processes to guide process-related actions [28]. The related common practices are as follows:

- "Decision Making Processes" are well defined which address "who" can do "what" and "who" has the "responsibility" for allocating "which" resources when anticipated and un-anticipated circumstances arise.
- It is important to have clearly defined "Roles" and "Responsibilities" within your BPM initiative (e.g. Business Analyst, Project Manager Etc.) including the associated reporting structures.
- Existence of processes for collecting the metrics which assist in measuring the alignment of business strategy with process output.
- Defining clearly the process management standards regarding process measures, issue resolution, rewards, etc., and including process coordination and initiatives should be clearly defined.
- BPM initiative is made regarding compliance, quality and standards as related to BPM governance.

2.1.6 Strategic Alignment

Strategic alignment is defined as the tight linkage of organizational priorities and enterprise processes, enabling continual and effective action to improve business performance [27]. The related common practices are as follows:

- Implementation a process improvement plan which provides information relating to the targets of the BPM Project and includes review and monitoring processes within the Plan.
- Having a clear and obvious collaborative link between "business strategy" and "business process design".
- Process designs are based around an enterprise process architecture where all Lines-of-Business are considered and that the process design and its value are available and accessible to the whole organization for consideration.
- For the process designs, there are well defined process outputs and Key Performance indicators in order that the process can be measured in terms of achieving a strategic goal or cost, time, flexibility, reliability etc.

- The strategy of process design is aligned to the priorities of individuals such as key stakeholders, LOB managers, shareholders, and government bodies etc.

Every organization has a different starting point and, as a result, different needs. Some already have a defined process; others are not as well developed. Some want to emphasize automation of the process, whereas others need better traceability, visibility and performance measurement. Either way, the first objective is to benchmark according to the best practices in different processes to deliver the most value [29]. Therefore, this paper applies FAHP to identify the best practices for core BPM elements and their related practices. In the following chapter, we introduce FAHP method, which is further used for the purpose of this research.

2.2 Analytical Hierarchy Process

AHP is applied in solving complex problem solving decision by using qualitative data. It's mostly used to assess the value and weight of intangibles. The basic idea behind AHP is that the goal components have to be identified it includes what re the integral parts and what the criteria are [30]. In the hierarchical format it has to contain a specific goal with its own criteria and also alternative criteria. After constituting the hierarchical structure that contained aim, criteria and alternative levels, the people in charge of making the decision are asked to have a pairwise comparison between criteria that is done in metrics. The judgment is done from the direct upper level criterion angle. Scoring of the comparison is performed in a relative basis. The importance and weight of each of the criteria are compared to that of another [31]. In the AHP method, interviews or questionnaires can be used to get importance weights of the components of the objective from decision maker's perspective.

2.3 Fuzzy AHP

As mentioned earlier, intangible values are often assessed by conventional AHP. However, conventional Saaty's AHP does not completely indicate the significance of qualitative criteria. The reason is that human uncertain thoughts cannot be reflected in the AHP scale [32]. In order to solve this problem, triangular fuzzy numbers (TFN) and conventional AHP are joined together to form Fuzzy AHP to remove the negative aspect of subjective assessments of decision makings [33].

Linguistic judgments are converted into TFNs (each set of TFN is indicted with i, j, k) which is structured in fuzzy pairwise comparison matrices. After these matrices are processed, the relative weights of all the elements and their corresponding ranking can be obtained. As mentioned earlier, there are many methods which are presented to develop the comparison metrics ([34], [35], [36]). As an example, according to Chang's method [35] which is one of the most popular methods, for each level of the constructed hierarchy, the pairwise linguistic judgments are converted in TFNs and structured in fuzzy comparison matrices. In this paper, the triangular fuzzy conversion scale is used to convert such linguistic scales into fuzzy scales in the evaluation model as shown in Table 2.

Table 2. Triangular fuzzy conversation scale [35]

Linguistic scale	Number put by respondents	Triangular fuzzy conversation scale	Triangular fuzzy reciprocal scale
EQUALLY important	1	(2/3, 1, 3/2)	(2/3, 1, 3/2)
WEAKLY MORE important	3	(1, 3/2, 2)	(1/2, 2/3, 1)
MODERATELY MORE important	5	(3/2, 2, 5/2)	(2/5, 1/2, 2/3)
STRONGLY MORE important	7	(2, 5/2, 3)	(1/3, 2/5, 1/2)
EXTREMELY MORE important	9	(5/2, 3, 7/2)	(2/7, 1/3, 2/5)

The calculation for the fuzzy weights according to Chang method is described in details as follows:

Matrix equation (1)

$$\tilde{A} = (\tilde{a}_{ij})_{n\times n} = \begin{bmatrix} (1,1,1) & \cdots & (l_{12}, m_{12}, u_{12}) & \cdots & (l_{1n}, m_{1n}, u_{1n}) \\ (l_{21}, m_{21}, u_{21}) & \cdots & (1,1,1) & \cdots & (l_{2n}, m_{2n}, u_{2n}) \\ \vdots & & \vdots & & \vdots \\ (l_{n1}, m_{n1}, u_{n1}) & \cdots & (l_{n2}, m_{n2}, u_{n2}) & \cdots & (1,1,1) \end{bmatrix}$$

Equation (2)

$$\tilde{a}_{ij} = (l_{ij}, m_{ij}, u_{ij}) = \tilde{a}_{ji}^{-1} = \left(\frac{1}{u_{ji}}, \frac{1}{m_{ji}}, \frac{1}{l_{ji}}\right) \quad i, j = 1, \ldots n;\ i \neq j$$

Represents the linguistic judgment for the items i and j; thus $\tilde{A}$ is a square and symmetrical matrix. For each row of $\tilde{A}$ it is possible to calculate the relative row sum as:

Equation (3)

$$\widetilde{RS}_i = \sum_{j=1}^{n} \tilde{a}_{ij} = \left(\sum_{j=1}^{n} l_{ij}, \sum_{j=1}^{n} m_{ij}, \sum_{j=1}^{n} u_{ij}\right) \quad i = 1, \ldots, n$$

By modifying the Chang's normalization formula using Wang and Elhag [37] correction, it is possible to obtain the normalized row sum $\tilde{S}_i$ as:

Equation (4)

$$\begin{aligned}\tilde{S}_i &= \frac{\widetilde{RS}_i}{\sum_{j=1}^{n}\widetilde{RS}_j} \\ &= \left(\frac{\sum_{j=1}^{n} l_{ij}}{\sum_{j=1}^{n} l_{ij} + \sum_{k=1,k\neq i}^{n}\sum_{j=1}^{n} u_{kj}}, \frac{\sum_{j=1}^{n} m_{ij}}{\sum_{k=1,k\neq i}^{n}\sum_{j=1}^{n} m_{kj}}, \frac{\sum_{j=1}^{n} u_{ij}}{\sum_{j=1}^{n} u_{ij} + \sum_{k=1,k\neq i}^{n}\sum_{j=1}^{n} l_{kj}}\right) \\ &= (l_i, m_i, u_i) \quad i = 1, \ldots, n\end{aligned}$$

The normalized row sums $\tilde{S}_i$ $(i = 1, \ldots, n)$ are then compared using the degree of possibility.

Equation (5)

$$V(\tilde{S}_i \geqslant \tilde{S}_j) = \begin{cases} 1 & \text{if } m_i \geqslant m_j \\ \frac{u_i - l_j}{(u_i - m_i) + (m_j - l_j)} & \text{if } l_j \leqslant u_i \; i, j = 1, \ldots, n; \; j \neq i \\ 0 & otherwise \end{cases}$$

Finally, the relative crisp weight of each item *i*, is calculated normalizing the degree of possibility values:

Equation (6)

$$w_i = \frac{V(\tilde{S}_i \geqslant \tilde{S}_j | j = 1, \ldots n; \; j \neq i)}{\sum_{k=1}^{n} V(\tilde{S}_k \geqslant \tilde{S}_j | j = 1, \ldots n; \; j \neq k)}, \quad i = 1, \ldots, n$$

Equation (7)

$$V(\tilde{S}_i \geqslant \tilde{S}_j | j = 1, \ldots n; j \neq i) = \min_{j \in \{1, \ldots, n\}, j \neq i} V(\tilde{S}_i \geqslant \tilde{S}_j) \quad i = 1, \ldots, n$$

The AHP methodology has been widely utilized in various fields: selection of a certain product, critical factors of BPR and six sigma, economic and management problem solving and so on. This method is specially used for intangibles evaluation as a method to give value weights to such parameters. It has been wide applied in prioritizing among various choices to arrive at a consensus decision [38]. This paper contributes to the use of FAHP to identify the best practices among the selected ones in this paper. The next chapter deals with applying this method to 30 practices for 6 BPM core elements. Having mentioned above references, we could not find specific references addressing application of this method to BPM area. However, there have been several studies that focused on identification of best practices in Business Process Redesign through case studies [39, 40]. The emphasis of the BPM references has been mostly on measurement and evaluation of business processes. For the remainder of this paper, we show the application of FAHP method in the following chapter.

3 Methodology

In our research, we are interested in developing a methodology for guiding and supporting the BPM execution and improvement by means of identification of best practices. In the first step, we have identified the most common BPM elements which are: information technology, culture, strategic alignment, methods, people and governance [14-16]. In the next step we found the common practices related to these core elements which were obtained through literature survey and interviews with process architects from BPM department in the selected three telecommunication companies in Denmark. The last step was to identify the best practices. For this purpose, we have again contacted process architects to fill out the questionnaire prepared for FAHP. We had received 45 responds. The relative importance (fuzzy weight) for each element and practice was determined by FAHP. The FAHP is used extensively in organizations that have carefully investigated its theoretical underpinnings. It is one of the most applied structured techniques that has produced extensive results in the following areas: ranking prioritization, resource allocation, benchmarking, quality management and conflict resolution [41]. This research also contributes to application of FAHP in prioritizing the BPM core elements and practices. The elements and the related practices are paired compared by each of the process architects based on qualitative scales given in table 2. Below small part of the questionnaire is indicated.

Table 3. Pair wise comparison of the core elements in FAHP questionnaire

BPM element	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	BPM element
Culture							X											Governance
Culture													X					IT

According to the Triangular fuzzy conversation scale shown in table 2, application of TFN works in a way that if a respondent chooses that "culture" is weakly more important (number shown is "3" in table 2) than "governance", in the calculation, the scale will be (1, 3/2, 2), so while comparing governance to culture, the scale will be the second column, which is (1/2, 2/3, 1). Therefore, for evaluating the opposite side, the reciprocal scale is applied to complete the analysis. Due to the space limit the calculation steps for the identification of the most important element among 6 core elements in implementing BPM are attached in the appendix in the following link: http://www.slideshare.net/slideshow/embed_code/36476418

The calculation indicates that among the 6 core elements of BPM, IT has obtained the highest importance in improving BPM application.

4 Results

The final results of the BPM practices that obtained the highest value weights among the 30 practices, for each practice and company are listed in table 4. Some practices have obtained the same value for two of the companies like “Documentation of methods for identifying and modeling "as is" and "to be" processed”. The value weights of all practices are avoided in this paper due to space limit.

Table 4. BPM practices with highest value weights

Element	Company A	Company B	Company C
Strategic alignment	For the process designs, there are well defined process outputs and Key Performance Indicators in order that the process can be measured in terms of achieving a strategic goal or cost, time, flexibility, reliability etc.	The strategy of process design is aligned to the priorities of individuals such as key stakeholders, LOB managers, shareholders, and government bodies etc.	Having a clear and obvious collaborative link between "business strategy" and "business process design.
Governance	Existence of processes for collecting the metrics which assist in measuring the alignment of business strategy with process output.	BPM initiative is made regarding compliance, quality and standards as related to BPM governance.	"Decision Making Processes” are well defined which address "who" can do "what" and "who" has the "responsibility" for allocating "which" resources when anticipated and un-anticipated circumstances arise.
Method	Documentation of methods for identifying and modeling "as is" and "to be" processes.	Documentation of methods for identifying and modeling "as is" and "to be" processes.	Existence of methods in place relating to process implementation and execution.

Table 4. (*continued*)

IT	There are BPM project management software tools available to facilitate the BPM initiative.	There are BPM software tools available supporting the process modeling, simulation, analysis and reporting	There are BPM software tools available supporting the process modeling, simulation, analysis and reporting
People	The willingness of individuals is within the BPM initiative to lead, to take responsibility and accountability for business processes encouraged and managed within the organization.	The willingness of individuals is within the BPM initiative to lead, to take responsibility and accountability for business processes encouraged and managed within the organization.	There are BPM process skills and expertise associated with the individuals Role and Responsibility.
Culture	There is a common belief within the organization of the value of BPM and its ability to benefit the business, and is that belief promoted well.	There is a managed active commitment from senior executives towards the BPM initiative.	There is a managed active commitment from senior executives towards the BPM initiative.

In summary, we provided through this work an insight into the contribution of FAHP to identify which BPM elements and practices are most important regarding BPM application development. Among the six core elements, the final result indicates that IT has the most important part in BPM implementation and improvement. Governance has obtained the least value weight in BPM application improvement. Similarly, each of the practices has obtained a value weight according to the data from questionnaire analyzed by FAHP. The practices which have obtained the highest value weights are categorized in table 4. These practices are thought to be considered with priority by BPM practitioners in relation to BPM application improvement.

5 Conclusion

In this paper, we have delivered the results of a study amongst process architects in three telecommunication companies in Denmark. The aim was to identify the best practices among the various practices related to six core elements of BPM including IT, culture, strategic alignment, methods, people and governance through the FAHP. The FAHP is applied instead of the AHP method in order to tolerate vagueness of the human uncertain thoughts. The 30 BPM practices and 6 core elements were analyzed by FAHP using a questionnaire. IT has gotten the highest fuzzy weight among the 6 core elements in relation to BPM implementation improvement. The calculation and the fuzzy weights are indicated in the appendix link mentioned earlier. The practices that obtained the highest value weights are shown in table 4.

We believe that in presenting the best practices in this paper, we provide support to the practitioner of BPM dealing with the mechanics of the process. The best practices may be used as a checklist for BPM application improvement and benchmarking. However, this research is not without weakness. One of the gaps is that, this research focused on telecommunication companies, so the results cannot be generalized for different industries as different organizations have different processes and different needs as well. Future research will focus on considering different industries and show the differentiation between the results obtained. The measurement of the impact of each practice on organizational performance including time, cost and quality is another step toward a more comprehensive research.

References

1. Lee, R., Dale, B.: Business process management: a review and evaluation. Business Process Management Journal, 214–225 (1998)
2. Frei, F.X., et al.: Process variation as a determinant of bank performance: evidence from the retail banking study. Management Science 45(9), 1210–1220 (1999)
3. Lockamy III, A., McCormack, K.: The development of a supply chain management process maturity model using the concepts of business process orientation. Supply Chain Management: An International Journal 9(4), 272–278 (2004)
4. Bosilj-Vukšic, V., Indihar-Štemberger, M.: The impact of business process orientation on financial and non-financial performance. Business Process Management Journal 14(5), 738–754 (2008)
5. Hammer, M., Champy, J.: Reengineering the Corporation: Manifesto for Business Revolution, A. Harper Collins e-books (2009)
6. Motwani, J., Kumar, A., Antony, J.: A business process change framework for examining the implementation of Six Sigma: a case study of Dow Chemicals. The TQM Magazine 16(4), 273–283 (2004)
7. Bach, N., Biemann, T.: Geschäftsprozessmanagement in Deutschland. Geschäftsprozessmanagement inside, pp. 1–27 (2004)
8. Forsberg, T., Nilsson, L., Antoni, M.: Process orientation: the Swedish experience. Total Quality Management 10(4-5), 540–547 (1999)
9. Frei, F.X., Harker, P.T.: Measuring aggregate process performance using AHP. European Journal of Operational Research 116(2), 436–442 (1999)
10. Zairi, M.: Business process management: a boundaryless approach to modern competitiveness. Business Process Management Journal 3(1), 64–80 (1997)
11. Johnson, M., Gustafsson, A.: Improving customer satisfaction, loyalty and profit: An integrated measurement and management system. Wiley. com (2006)
12. DeToro, I., McCabe, T.: How to Stay Flexible and Elude Fads: With business process management, organizations will avoid total allegiance to one approach. Quality Progress 30, 55–62 (1997)
13. Silvestro, R., Westley, C.: Challenging the paradigm of the process enterprise: a case-study analysis of BPR implementation. Omega 30(3), 215–225 (2002)
14. Rosemann, M., de Bruin, T.: Towards a Business Process Management Maturity Model. In: ECIS (2005)
15. De Bruin, T.: Business process management: theory on progression and maturity (2009)
16. Jeston, J., Nelis, J.: Business process management. Routledge (2014)

17. McCaffrey, J.: Test Run-The analytic hierarchy process. MSDN Magazine, 139–146 (2005)
18. Toro, I.D., McCabe, T.: How to stay flexible and elude fads. Quality Progress 30(3), 55–62 (1997)
19. Pritchard, J.-P., Armistead, C.: Business process management–lessons from European business. Business Process Management Journal 5(1), 10–35 (1999)
20. Spanyi, A.: Business Process Management (BPM) is a Team Sport: Play it to Win. Meghan Kiffer Pr (2003)
21. Vom Brocke, J., Rosemann, M.: Handbook on business process management 1. Springer (2010)
22. Trkman, P.: The critical success factors of business process management. International Journal of Information Management 30(2), 125–134 (2010)
23. Žabjek, D., Kovacic, A., Štemberger, M.I.: The influence of business process management and some other CSFs on successful ERP implementation. Business Process Management Journal 15(4), 588–608 (2009)
24. Attaran, M.: Exploring the relationship between information technology and business process reengineering. Information & Management 41(5), 585–596 (2004)
25. Burlton, R.: Delivering business strategy through process management. In: Handbook on Business Process Management 2, pp. 5–37. Springer (2010)
26. Ravesteyn, P., Batenburg, R.: Cultural differences in implementing business process management systems (2010)
27. Rosemann, M., vom Brocke, J.: The six core elements of business process management. In: Handbook on Business Process Management 1, pp. 107–122. Springer (2010)
28. Ngai, E.W., Law, C.C., Wat, F.K.: Examining the critical success factors in the adoption of enterprise resource planning. Computers in Industry 59(6), 548–564 (2008)
29. Drew, S.A.: From knowledge to action: the impact of benchmarking on organizational performance. Long Range Planning 30(3), 325–441 (1997)
30. Zahedi, F.: The analytic hierarchy process—a survey of the method and its applications. Interfaces 16(4), 96–108 (1986)
31. Saaty, T.L.: Decision-making with the AHP: Why is the principal eigenvector necessary. European Journal of Operational Research 145(1), 85–91 (2003)
32. Özdağoğlu, A., Özdağoğlu, G.: Comparison of AHP and fuzzy AHP for the multi-criteria decision making processes with linguistic evaluations. İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi 6(11), 65–85 (2007)
33. Kwong, C., Bai, H.: Determining the importance weights for the customer requirements in QFD using a fuzzy AHP with an extent analysis approach. Iie Transactions 35(7), 619–626 (2003)
34. Buckley, R.: Distinguishing the effects of area and habitat type on island plant species richness by separating floristic elements and substrate types and controlling for island isolation. Journal of Biogeography, 527–535 (1985)
35. Chang, D.-Y.: Applications of the extent analysis method on fuzzy AHP. European Journal of Operational Research 95(3), 649–655 (1996)
36. Csutora, R., Buckley, J.J.: Fuzzy hierarchical analysis: the Lambda-Max method. Fuzzy Sets and Systems 120(2), 181–195 (2001)

37. Wang, Y.-M., Elhag, T.: Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment. Expert Systems with Applications 31(2), 309–319 (2006)
38. Calabrese, A., Costa, R., Menichini, T.: Using Fuzzy AHP to manage Intellectual Capital assets: An application to the ICT service industry. Expert Systems with Applications (2012)
39. Reijers, H.A., Liman Mansar, S.: Best practices in business process redesign: an overview and qualitative evaluation of successful redesign heuristics. Omega 33(4), 283–306 (2005)
40. Mansar, S.L., Reijers, H.A.: Best practices in business process redesign: use and impact. Business Process Management Journal 13(2), 193–213 (2007)
41. Forman, E.H., Gass, S.I.: The analytic hierarchy process-an exposition. Operations Research 49(4), 469–486 (2001)

Customer Oriented Management of Changes in ERP Systems: The Vendor's Side

Erika Asnina, Mara Pudane, Marika Svike, and Gundars Alksnis

Institute of Applied Computer Systems, Riga Technical University
Meža iela 1 k-4, Riga, LV-1048, Latvia
{Erika.Asnina,Mara.Pudane,Gundars.Alksnis}@rtu.lv,
msvike@gmail.com

Abstract. The paper discusses a business case when customer's deployment of the Enterprise Resource Planning solution (ERP) is integrated with third-party's software. Customer's change management highly depends on vendor's information about changes between ERP versions. The problem is that the vendor cannot predict the impact of implemented changes on every customer's business. In the suggested conceptual customer-oriented solution, the conjoint assessment evaluates the size, the scope and the essentiality of ERP changes in order to provide data on changes for a particular customer. The resulting difference model could be enhanced with results of analysis of collected access statistics for customer's ERP packages. The solution uses a unified data format for data interchange. The final model of ERP changes and their essentiality for a particular customer could be visualized for the further application.

Keywords: change management, ERP, XMI, XML, access statistics.

1 Introduction

Enterprise Resource Planning systems (ERP) are widely-used (usually commercial) information systems (ISs) [1] which support customer's business. Evolution of IT technologies, legislation, customers' demands and business activities forces organizations to make the process of updating software more predictable and controlled by introducing change management (CM) activities [2], [3]. Very little attention has been paid to ERP maintenance questions before 2001 [3], and only small changes in this field can be observed at the present time.

A challenge arises when CM is needed for ERPs which are highly integrated with third-party ISs on the customer side. In this case, a customer owns an instance of a database (DB) and some functional parts (subsystems, modules, components, etc.) of ERP. On the basis of his own ERP instance, the customer builds business-specific functionality or uses an integrated software solution.

The business case we discuss in this paper states that the ERP vendor implements changes without negotiations with its customers and is not familiar or does not have complete access to the co-integrated ISs. On the other hand, the customer does not

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 320–334, 2014.

have complete access to ERP source code and cannot predict the impact of ERP changes on its ISs. Thus, both the vendor and the customer are not able to perform overall CM and reduce unpredictable failures in the functionality of customer's ISs.

The aim of the research is to propose the conceptual solution to this problem that could improve customer service on the vendor side by reducing unpredictability to customers. In [4] we have presented the initial discussion on the data format for this solution. Here, we present more complete results, i.e., the entire scheme of the solution as well as its validation results. The suggested general scheme of evaluating customer-specific essentiality of ERP changes can be added to already existing CM on the vendor's side in order to improve analysis of change requests (inner as well as external), to evaluate change impact on the customer's ISs, and to provide the results to the customer for further analysis on his side in a format suitable for both automated handling and manual analysis.

The business case is explained in detail in Section 2. Characteristics of changes, change impact analysis (CIA) methods, and visual representation of changes are discussed in Section 3. Data export, storing and interchange formats are discussed in Section 4. Section 5 presents the suggested conceptual solution, i.e., the evaluation method and the data format as well as results of validation. Section 6 discusses related work. Section 7 concludes the paper and discusses further research direction.

2 The Business Case, Open Questions and Research Methods

The ERP implementation is adapted to customer business needs and can be integrated with other ISs independently from the vendor. A lack of information about implemented changes raises the following questions we try to solve: How to assess an impact of changes implemented in ERP to customer's business? How to handle huge amounts of data (necessary for the assessment), which come from different sources and have different structures and types? We *do not* consider adoption of ERPs or implementation of customer-initiated changes in ERPs in this paper.

The business case has two sides, namely the vendor's side and the customers' side [4]. The ERP consists of *hundreds* of primitive tables and *thousands* of coded units (objects, functions, and reports). For example, by changing a table to a new version, corresponding objects are damaged affecting dependent units. If a vendor knew essentiality of the effect raised by the changes to other units, he could inform the customer faster and provide more complete and customer-specific characteristics of changes. On the other hand, if a customer knew the more important units for his business, he could plan to update them faster and to take more time to update less important ones. For this purpose we propose to develop a logical analytical component that would give answers to questions stated by the vendor [4]: How to characterize the implemented changes and their importance? How to compare two business models which are built from primitive tables and objects? How to assess the size and the scope of changes and their influence on the customer's ISs, which are integrated with the changed one? In what form the implemented changes are to be represented for change analysis conducted by people?

According to the questions, a conceptual scheme of the analysis block consisting of three tools for exporting and comparing models, collecting statistics and visualizing the resulting model of differences was created and confirmed with vendors [4]. It is based on the following four sub-goals and ten tasks (Fig.1):

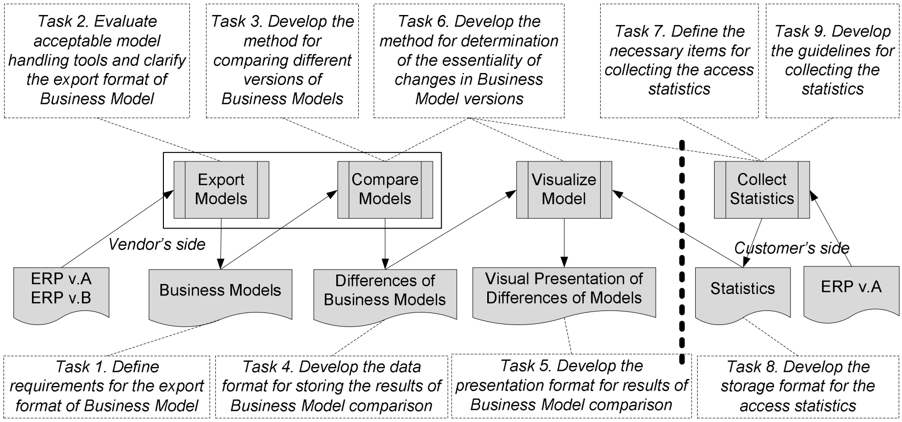


Fig. 1. Research objects and tasks

- *Sub-goal 1:* Determination of the size of implemented changes that helps the vendor in understanding the differences between any two versions of the ERP (ERP v.A and ERP v.B in Fig. 1). It includes *Task 1*, *Task 2*, *Task 3* and *Task 4*.
- *Sub-goal 2:* Determination of the scope of implemented changes that focuses vendor's and customer's attention mainly on customer-specific changes. It includes *Task 6*, *Task 7*, *Task 8*, and *Task 9*.
- *Sub-goal 3:* Presentation of the results in a human-readable visual form in order to make further analysis easier for a customer. It includes *Task 5*.
- *Sub-goal 4:* Validation of the results by prototyping. It includes the prototyping itself and evaluating the test results (*Task 10*).

3 Evaluation and Impact Analysis of Changes

3.1 Determination of Importance of Software Changes

Precise evaluation of importance of the change (*Task 6*) requires taking into account two components - business and software [5].

The business component allows evaluating the impact of changes on the organization's business processes. The criticality/importance of the change is determined by its non-quantitative characteristics - source and motivation [6], [7], [8], [9]. The sources could be found in stakeholders' requests or customers' demands, policy, legislation or government regulations [6], [7], [8], [9], resource constraints [6] as well as technological specifications [7], [8], [9]. The motivation of the change determines whether it is perfective, adaptive, corrective, or preventive [6]. In the two latter cases, the importance of the change should be higher. Expert-defined rules can be used [10] for evaluation of the non-quantitative characteristics.

The software component allows evaluating the impact of changes on the organization's applications. Changes can be distinguished by their quantitative characteristics: logical or physical location [11], transformation operations [12], [13], system properties affected by changes [6], the size, impact, structure, vocabulary, change scope, dependency and correlation [6], [12]; a period of time of last changes, and significance of the changes (the union of the location, size, impact, vocabulary, and change scope) [14].

3.2 Determination of the Size of Changes by Software Model Comparison

The size of the implemented changes (*Task 3*) can be determined by comparing models of the software structure (Section 4). Due to the ERP specifics, the main focus is put on object-oriented models but not on data, information or process flows. The discussion about model comparison includes two directions, namely similarity search [15] and control of model versions in repositories [16]. Methods of model comparison can be divided into two groups: high-level abstraction methods, which provide an algorithm or framework for any type model comparison, and language or notation-specific methods, e.g., a general rule-based framework [17], comparison algorithms for UML (Unified Modeling Language) diagrams [18], XML (eXtensible Markup Language) documents [19], XMI (XML Metadata Interchange) documents [20], [21], and XMI usage as an UML representation for model comparison [22].

Table 1. Enumeration of tool characteristics. Denotation: *V* - visualization, *C* - comparison.

Tool	Status	V/C	Export Format	Models
Altova	Commercial	V/C	XMI	UML, ERD
Sparx Enterprise Architect	Commercial	V/C	XMI	UML
IBM System Architect	Commercial	V/C	XML Schema	UML
IBM Rational Rose	Commercial	V/C	XMI	UML, ERD
Toad Data Modeler	Commercial	V/C	XMI, ORDB	UML, ERD
Visual Paradigm	Commercial	V/C	XML Schema, ORDB	UML
Protégé	Free	V	XML Schema	-
Eclipse Modeling Framework	Open source	V/C	XMI	UML
EMF Papyrus	Open source	C	XMI, RDB	UML

Model comparison is already automated (column 3 in Table 1), e.g. in Altova [23] and IBM Rational System Architect [24]. As a latest trend, a question on how to compare models due to the semantics raises [25]. The results of the comparison can be visualized [26].

3.3 Determination of the Scope of Changes by Impact Analysis Techniques

Change impact analysis (CIA) is a collection of techniques which allow understanding the potential impact (the scope) of software changes on other parts of the system (*Task 6*) [27]. The impact can be large even from small changes, and vice versa. The goal of the CIA is to identify the ripple effects and to predict the possible side effects of changes [27]. Two categories of CIA techniques exist [9]: implementation-based CIA and model-based CIA. The both categories require completely accessible or known software source/execution code or models. This limits their use since third-party ISs are not accessible in our business case.

Another opportunity is to collect statistical data (*Tasks 7, 9)* on access to software or databases (DBs). Statistics [28] helps in software maintenance. In order to define a frequency of accesses to labeled modules the following data can be collected: the data on DB module access rates [29], the number of accesses to tables and indexes (disk blocks, individual rows), user functions, server activities [30], and function invocations during program execution and compilation [31]. As mentioned in [29], [30], [31], a collector of statistical data can be a module, a subsystem, an analysis tool, or logging functional code. The placement of the collector can be local (in the initialization methods, concrete methods or routines, modules, files) or global on the server side (as a subsystem of the DB management system (DBMS) or an independent program). The collected data can be stored locally in log/trace files. The statistical data can be analyzed separately for each module, and then final analysis can be done in a centralized way by using a set of the intermediary results [29], completely in a centralized way by means of the DBMS [30], or in the tracking DB [31].

To sum up, the main problem can be an increase of system workload during the query and function execution. It can be solved by collecting and partially analyzing data locally, and, when software finishes its work, transmitting the statistical data files to the central storage for analysis.

3.4 Visual Presentation of Models and Changes

Software changes can be represented as a set of difference elements (*Task 6)*. Their visualization could help vendors and customers in change comprehension, change prediction, and contribution analysis [32]. The visualization model [11], a set of invariants of the mapping from the abstract data to visual objects, helps with getting a simplified view on the changes and facilitates understanding the change characteristics. The visualization recognizes patterns among the changes (e.g., feature removals, method calls replacements), and other aspects such as complexity or semantic impact of the changes [12].

Tools and systems [6], [11], [33], [34], [12], as well as frameworks and models [34], [11], [14], [35], [13] are used for change visualization. The tools can present the extracted and derived information in a form of text [11], [13], [33], [36], hypertext or graphics [34], a combination of text, hypertext or graphics [13], [37], [34], different kinds of views (statistical, graph, special) [34], [13], [37], graph drawings [14], and as annotated source code and control flow graphs [35]. Integration of techniques and combination of different visualization forms, e.g., visual variables such as color, position, size, and transparency, animation or motion, as well as user- or tool-generated abstractions [34], can support customer's analysis more effectively. Due to the very large amount of data on ERP, we suggest using tools that can automate the visualization process.

4 Data Format for Export, Storage and Interchange

Specifics of the business case is the need to operate with a great amount of data of different types, namely large software models, models of differences between versions, statistics collected during execution of ERP and third-party software by customers. The business case requires not only automated comparison of ERP versions and evaluation of the scope of implemented changes, but also further manual analysis of the results. Therefore, data export, storage and interchange format must be suitable for automated data handling and further visualization in human-friendly forms. The results of *Tasks 1, 2, 4, 5,* and *8* are summarized here.

The natural way to reflect the ERP structure is the use of models. An IS model is a set of diagrams which show systems structure from different viewpoints and behavior in different scenarios [38] in some modeling languages [39], [40]. The languages vary from simple structural like Entity-Relationship Diagrams (ERD) [41] to complex ones like object-oriented UML [40], [42], where levels of abstraction depend on a domain [41]. Main benefits of a model usage in CM are flexibility to add and simulate components which could be too expensive in reality, possibility to involve system analysts and other stakeholders, and reverse engineering [38], [43], [44].

Metadata Model Storage (Tasks 1, 2, 4). Software models contain system metadata, e.g., DB tables, classes with their properties, action scenarios, etc. The result of *Task 1* are requirements defined for the export format of software models: It should be object-oriented in order to specify objects, inheritance, aggregations, and relations "one-to-many"; standardized - to be imported/exported with existing tools; and flexible - to be able to add supplementary parts, e.g., access statistics.

There are several ways how to store model metadata, e. g., in text files using XML Schema and XML [45], Rich Text Format (RTF) [46], in (object-) relational DB - (O)RDB tables as columns [47], or a combined approach [43]. Only a few XML-based metadata storage templates for object-oriented models which can be considered as metadata storage methods, exist, such as IEEE LOMv1.0 for learning objects [48] or XMI [42]. XML is a de facto standard for storing data in text files [49]. Its main characteristics are self-reflexive, human and machine readable, extensible, flexible, platform-neutral, portable, specially designed to carry data, facilitate logical data management, handle only content (while presentation is addressed separately) and support validation of external data [50], [51], [52]. The structure of XML files may vary from a flat regular data-centric structure to a deep irregular document-centric structure. Thus, it could be used for the representation of all kinds of data [53]. XML and XMI advantages are expressive representation power of relational data, report formats and semi-structured documents [50], combination of data and metadata that enables smart searches and data interchange [50], achievement of content integration, intelligence and reuse, support of storage and manipulation with metadata [52].

After evaluation of method characteristics [4] and tool analysis (columns 4 and 5 in Table 1), it is found that the XMI format corresponds more to the required characteristics than others; although it requires improvements of flexibility. Still, it is based on the XML Schema that has the required flexibility. Thus, as a result of *Tasks 1, 2* and *4,* both XMI and XML Schema have been selected as suitable data formats.

Storing, Retrieving and Querying Statistics (Task 8). Storage and analysis of statistics in XML files can be supported by XML-enabled DBs (XED), document-based (NoSQL) and native XML-based DBs (NXD). These activities have three aspects: storage of the collected data, the efficiency of exporting collected statistics to the XML-based format, and the efficiency of querying both statistics and models.

Table 2. XML-enabled (XED), document-based (NoSQL) and native XML-based (NXD) database performance in handling XML documents. NXDs showed the better results, but they also require well-thought-out design of the XML structure [4].

Characteristics	XEDs	NoSQL	NXDs
efficiency in XML data preparation	low	high	high
efficiency in querying XML files	low	high	high
efficient XML structure	simple; flat	complex; flat	complex; *flat*
efficiency in modifying XML files	lower than NXD	lower than NXD	high
impact of an XML file storage way	have impact	have impact	*have impact*
scalability	high	high	high
extensibility	high	low	high
portability	high	low	high

The best efficiency of handling XML documents showed NXDs (Table 2), but they also require well-thought-out “flat” design of XML structures as well as a method of storage of XML documents. The explanations are as follows:

- XML data preparation, querying, modification, and recommended structure: the “flat” design of the XML document structure is the most efficient for all DBs, but XEDs require them to have a simple structure [49]. The reason is that XEDs require manual definition of mapping rules between XML document structures and 3NF [51], [52], NoSQL uses key-value stores, BigTable implementation, document stores, and graph DBs [51], but NXDs can store XML raw data and documents as a single storage unit [51] and allow data to be stored, queried, combined, secured, indexed in order to enhance query performance, etc. [50]. The XED querying mechanism is suitable for XML documents with simple structures, since requires conversion to SQL codes [51], NoSQLs and NXDs are efficient due to XQuery technology maturity [51], [49]. But for complex structures the latter two may require indexing systems [53], [49]. NDXs operate faster than XEDs when complex elements are inserted and whole documents are inserted or deleted, or data are appended to the end of files [49]. However, the way of storing files (as one big file or a number of files) do have impact on the performance of NDXs [49].
- Scalability, extensibility, and portability: XEDs can accept more than 1 GB of data, NoSQL approaches have no limit of the amount of data, but have potential DB administration and maintenance problems [51]; in NXDs the upper bound of the

file size is about 9 MB per file [49]. In XEDs extensibility and portability is high due to the mature technology used here [52], the same is true for NXDs due to high interoperability with other software systems with different data structures; however, NoSQL loses in short terms here [51].

Visualization of XML documents (Task 5). XML documents can be visualized as a tree structure, while XMI documents foresee also graphical representation. To visualize model changes (three subclasses add, replace and modify of a difference class [54]) and to transform XML/XMI structures into a hypertext document two wide-used technologies exist – XSLT [55] and CSS [56], [57]. XSLT (eXtensible Stylesheet Language Transformations) can transform XML in formatted HTML document [55], but the CSS (Cascading Style Sheet) style can handle information of element style, color and fonts. XSLT and XQuery can be used together.

5 The Conceptual Solution: The Method and Data Format

The requirements for the metadata exchange format state that it should be standardized in order to be shared between different tools, flexible enough to include information of change impact on a customer and visualization parameters, and object-oriented due to the nature of the business case. The XMI technology was found to be the most compliant to the requirements (Table 1). The additional XMI benefits are the ability to show changes between two XMI documents that is an implementation of CM directly in the format notation, as well as the possibility to handle XMI documents as such and together with general XML documents.

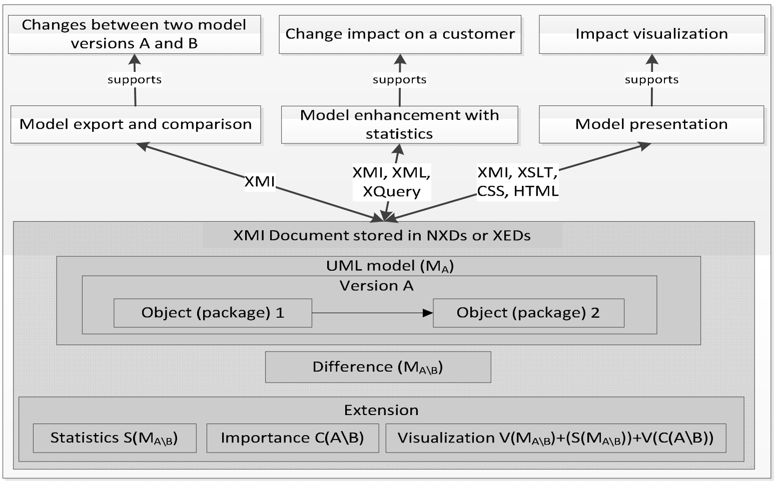


Fig. 2. The conceptual solution for the data format and technologies used

The XMI document structure (Fig. 2) consists of the pre-defined XMI and UML components (Documentation and object definitions) and the XMI Extension part that should contain the following four components: difference between two XMI files

$M_{A\backslash B}$; a complex element for statistics of changed elements for the customers $S(M_{A\backslash B})$; characteristics of change importance $C(A\backslash B)$; and visualization of both models themselves and change essentiality $V(M_{A\backslash B})+V(S(M_{A\backslash B}))+V(C(A\backslash B))$. In all denotations A and B are two versions of the ERP. The chosen XMI format allows model interchange and modifications (adding and removing the extension parts) between tools and a data storage. The literature survey showed that NXDs and XEDs could be used as databases, while NXDs have more benefits (Table 2).

Fig. 3 illustrates how XMI/XML metadata and data documents are used in the developed method for evaluation of changes between two versions of ERP Business Model. This method allows achieving the main research goal. It consists of five steps, which can be automated as XQuery and XSLT functions, except model comparison that already is automated in different tools (Table 1).

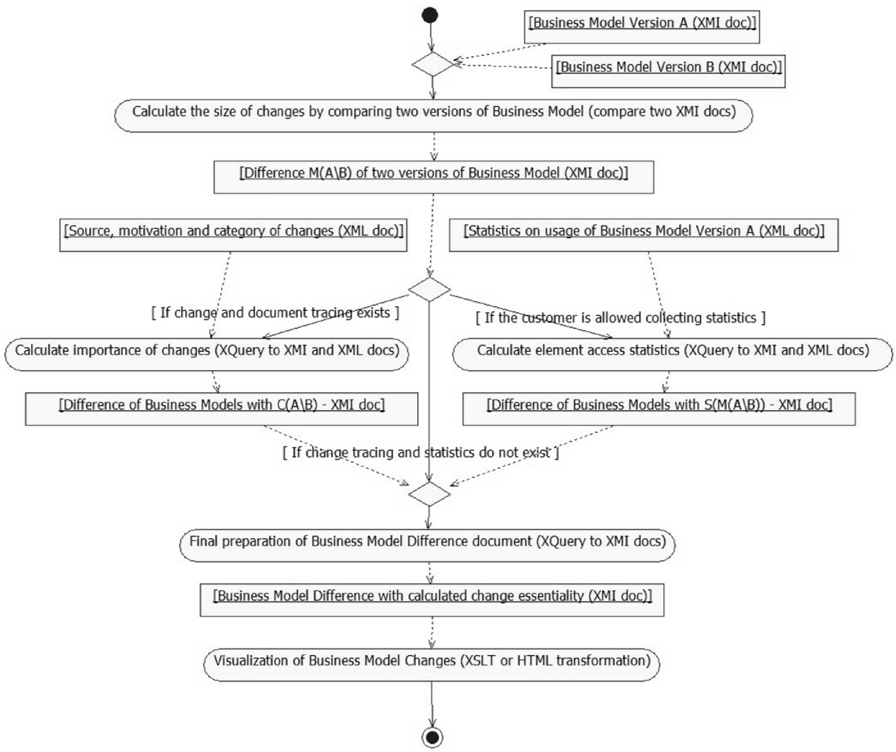


Fig. 3. The method for evaluation of changes between two versions of Business Model

Step 1 "Calculate the size of changes by comparing two versions of Business Model" uses two versions A and B of ERP Business Model, exported as two XMI documents (M_A and M_B). Version A is the original version of ERP Business Model, but version B contains metadata of ERP Business Model with introduced changes. Differences between two models $M_{A\backslash B}$ (or in other words, the size of changes) can be obtained by using a model comparison method and saved as a separate XMI document (Section 3.2).

Step 2 "Calculate importance of changes" is executed, if an organization has change and document tracing, i.e., it is possible to identify sources, motivations and categories of changes from existing change requests and management documents. The result is an XMI document with differences $M_{A\backslash B}$ and importance of introduced changes C(A\B). The calculation is based on the following principles (Section 3.1):

1. Identify introduced changes *changeRef* for each element in differences model $M_{A\backslash B.}$
2. Calculate the common importance */changeImportance* for changes in *changeRef*:

- Source: If the source type of a change is any of *ProductEvolutionPolicy*, *Legacy*, *Regulators*, *CompetitiveForces*, *TechnologyEvolution*, then the value is "*high*"; if *ClientDemands*, then "*medium*"; If *SocialEnvironment*, then "*low*";
- Motivation: If the motivation type is *DefectCorrection*, then the value is "*high*"; If *FunctionalEnhancement*, then "*medium*";
- Change category: If the change category type is *Corrective* or *Preventive*, then the value is "*high*"; If the change category type is *Perfective* or *Adaptive*, then the value is "*medium*";
- The value of the common importance of the change source, motivation and category is equal to the highest calculated value.

3. Assign the calculated value to each changed element.

Step 3 "Calculate element access statistics" is executed in order to calculate general and proportional usage frequency for each element (Section 3.3):

1. Calculate usage frequency for each element per each registered program, application, interface in collected statistics, as well as the summarized value */accessCount*;
2. Calculate usage frequency */generalChangeScope* of all elements.
3. Calculate the proportional value of usage frequency for each element */accessProportionalValue* = (*/accessCount* * 100) / (*/generalChangeScope*).

The result is an XMI document with differences $M_{A\backslash B}$ and calculated access statistics $S(M_{A\backslash B})$.

Step 4 "Final preparation of Business Model Difference document" is dedicated for joining differences, change importance and access statistics in one XMI document and sorting them by change essentiality. The higher is importance and/or the proportional access value of the element, the higher is change essentiality of the element. According to the selected approach, supplementary parts (importance, statistics) can be joined in one document or stored separately, by keeping references to the main document, namely the document of Business Model Difference.

Step 5 "Visualization of Business Model Changes" transforms the XMI document into a human-friendly form by using XSLT or HTML (Section 3.4).

The validity and workability of the proposed solution were verified by prototyping by using existing tools. The prototyping activities included five tasks: verification of the model export format, model comparison, work with the access statistics and evaluation of importance of changes, as well as visualization of model differences with and without customer-specific information. The tool prototype has been developed for model export. The exported model represented a part of the real ERP, which consisted of more than 200 objects. Other functionality was provided by already exiting and available software tools: EMF Compare, <oXygen/> XML Developer 15.2, Altova XMLSpy 2014 Enterprise Edition, Stylus Studio X14, and Altova StyleVision 2014. The results showed that the proposed method and data format are workable and allow achieving the goal. However, selection of already existing tools requires special attention for the ability to read/import XMI and UML metamodels, while the custom XML schemas are well understood by tools.

6 Related Work

Although orientation on customer needs is recognized in ERP adoption, it is expressed quite differently. Authors in [58] concentrate on SAP GUI (Graphical User Interface) sustainability due to IT innovations, by separating business logic and presentations. They do not consider importance of changes to customer's third-party software. The necessity for organizations to be responsive to internal and external changes to become competitive is also noted in [2], where authors focus on internal adoption and institutionalization of ERP within the organization by means of organizational responsiveness, but they do not consider maintenance activities of already adopted ERPs. Authors in [3] do consider maintenance activities and give their taxonomy, but their research is based on two case studies where organizations maintain their own ERP systems. Their research showed that in most cases changes in ERP fall into enhancive (adaptive and perfective) and corrective categories. Authors proposed five fundamental business benefit categories of enhancement maintenance for facilitating the prioritization of tasks of change implementation. The same authors in [59] proposed a preliminary ERP maintenance model. The model is based on results of a concrete case study and is not applicable for the business case discussed in this paper, since it considers user-initiated changes in their own ERPs. Risk management also can be used as a supplementary activities in ERP maintenance (e.g. [60]), but in this business case it is rather the customers' side than the vendor's one, since the vendor has no access to the customer's business-specific details.

Other solutions are human-related and solve ERP change acceptance by organizations' employees to prevent their dysfunctional behavior (e.g. [61]) and cultural transformations during ERP adoption (e.g. [62]).

7 Conclusions

The business case discussed in this paper points to the lack of the way how the vendor can predict the impact of changes implemented in the ERP on his customers' business when the access to customer's integrated ISs is impossible. Based on the defined research questions and the review of related work in the field, the conceptual solution is proposed and suitable technologies are determined.

This solution is useful for ERP vendors, because it allows improving customer service as well as analyzing which ERP elements are most often used by customers and which ones are critical for the big customers. Despite that the given research is still in progress, the results presented here allow vendors to choose the most proper technology and to implement the solution – by using XMI and XML as main technologies for software model export, analysis, comparison, modification, as well as graphical and web-based presentation of the content; NXDs as the main technology for model storage and XQuery - for analysis and querying XMI/XML documents; as well as HTML, XSLT and CSS as technologies for presenting XMI documents in human-friendly forms.

Benefits for the customer are human-friendly presentation of customer-specific detailed information of the implemented changes. The customer can be supported with the visualized XMI model that contains the description of ERP elements and their relations, indicators on elements that are affected by changes and that are most important for the customer, and also indicators of importance of the changes from the business perspective. This solution does not require any additional investments from the customer side.

The solution was prototyped, nevertheless it still requires additional evaluation of the chain of supporting software tools. The tools must support UML and XMI metamodels as well as customized XML Schemas. It is planned to find or implement the supporting toolset and to validate the proposed solution and recommended technologies (XMI, NXDs, HTML, XSLT, and CSS) for the real scale of the ERP.

Acknowledgment. The research has been conducted within the framework of European Regional Development Fund's project "Information and communication technologies competence center" Nr. KC/2.1.2.1.1/10/01/001 (Contract No. L-KC-11-0003) activity 1.1 "Technologies of service configuration management" and activity 1.2. “Technologies for managing services in workflows”.

References

1. Columbus, L.: ERP Market Share Update: SAP Solidifies Market Leadership. Forbes (May 2013), `http://www.forbes.com/sites/louiscolumbus/2013/05/12/2013-erp-market-share-update-sap-solidifies-market-leadership/`
2. Pishdad, A., Haider, A.: Responding to external and internal changes: Organizational responsiveness pressures in institutionalising ERP systems. In: Proceedings of PICMET 2013: Technology Management in the IT-Driven Services (PICMET), July 28-August 1, pp. 2305–2312. IEEE (2013)
3. Ng, C.S.P., Chan, T., Gable, G.G.: A client-benefits oriented taxonomy of ERP maintenance. In: Proceedings of the IEEE International Conference on Software Maintenance 2001, pp. 528–537. IEEE (2001)
4. Asnina, E., Pudane, M., Svike, M., Alksnis, G.: Customer-Oriented Change Management for ERP Systems. In: Proceedings of the 15th International Conference on Computer Systems and Technologies - CompSysTech 2014, Ruse, Bulgaria, June 27, 2014. ACM, New York (2014) (accepted)

5. Asnina, E.: Essentiality of Changes in Business Models. In: Proceedings of the 2nd International Business and Systems Conference, BSC 2013, pp. 44–51. Riga Technical University, Riga (2013)
6. Williams, J.B., Carver, J.C.: Characterizing software architecture changes: A systematic review. Information and Software Technology 52(1), 31–51 (2010)
7. Osterwalder, A.: The Business Model Ontology: A Proposition. In: A Design Science Approach. PhD. thesis. Universite De Lausanne, Ecole Des Hautes Etudes Commerciales (2004)
8. de Reuver, M., Bouwman, H.: Business model dynamics: a case survey. Journal of Theoretical & Applied Electronic Commerce Research 4(1), 1–11 (2009)
9. Chen, C.-Y., Chen, P.-C.: A holistic approach to managing software change impact. Journal of Systems and Software 82(12), 2051–2067 (2009)
10. Jakimavicius, T., Kataria, P., Juric, R.: Semantic support for dynamic changes in enterprise business models. Transactions of the SDPS: Journal of Integrated Design and Process Science 14(2), 1–11 (2010)
11. Voinea, L., Telea, A.: Visual data mining and analysis of software repositories. Computers & Graphics 31, 410–428 (2007)
12. Gómez, V.U., Ducasse, S., D'Hondt, T.: Visually characterizing source code changes. Science of Computer Programming (2013), doi:`http://dx.doi.org/10.1016/j.scico.2013.08.002` (available online August 29, 2013)
13. Girschick, M., Darmstadt, T.: Difference Detection and Visualization in UML Class Diagrams. Technical Report TUD-CS-2006-5 (2006)
14. Collberg, C., Kobourov, S., Nagra, J.: A system for graph-based visualization of the evolution of software. In: Proceedings of the 2003 ACM Symposium on Software Visualization (SoftVis 2003), p. 77. ACM, New York (2003)
15. Dumas, M., Banuelos, L.G., Dijkman, R.: Similarity search of business process models. Bulletin of the Technical Committee on Data Engineering 32, 23–28 (2009)
16. Dijkman, R., la Rosa, M., Reijers, H.: Managing large collections of business process models—Current techniques. Computers in Industry, 91–97 (2012)
17. Kolovos, D.S., Paige, R.F., Polack, F.: Model Comparison: A Foundation for Model Composition and Model Transformation Testing. In: Proceedings of the 2006 International Workshop on Global Integrated Model Management (GaMMa 2006), pp. 13–20. ACM, New York (2006)
18. Kelte, U., Wehren, J., Niere, J.: A Generic Difference Algorithm for UML Models. Software Engineering 64, 105–116 (2005)
19. Cobena, G., Abiteboul, S., Marian, A.: Detecting changes in XML documents. In: Proceedings of 18th International Conference on Data Engineering, pp. 41-52. IEEE (2002)
20. Nentwich, C., Emmerich, W., Finkelstein, A., Ellmer, E.: Flexible consistency checking. ACM Trans. Softw. Eng. Methodol. 12(1), 28–63 (2003)
21. Nentwich, C., Capra, L., Emmerich, W., Finkelstein, A.: xlinkit: A Consistency Checking and Smart Link Generation Service. ACM Transactions on Internet Technology 2(2), 151–185 (2002)
22. Ohst, D., Welle, M., Kelter, U.: Differences between Versions of UML Diagrams. In: Proceedings of the 9th European Software Engineering Conference, pp. 227–236. ACM Press (2003)

23. Altova. Database Schema Comparison Tool, http://www.altova.com/databasespy/database-structure-compare-tool.html
24. Letkeman, K.: Comparing and merging UML models in IBM Rational Software Architect: Part 3, http://www.ibm.com/developerworks/rational/library/05/802_comp3/
25. Altmanninger, K.: Models in Conflict – Towards a Semantically Enhanced Version Control System for Models. In: Giese, H. (ed.) MoDELS 2007 Workshops. LNCS, vol. 5002, pp. 293–304. Springer, Heidelberg (2008)
26. Engelen, L., van den Brand, M.: Integrating Textual and Graphical Modelling Languages. Electron. Notes Theor. Comput. Sci. 253(7), 105–120 (2010)
27. Li, B., Sun, X., Leung, H., Zhang, S.: A survey of code-based change impact analysis techniques. Softw. Test. Verif. Reliab. 23, 613–646 (2013)
28. Dodge, Y.: The Oxford Dictionary of Statistical Terms, OUP (2006)
29. Dan Levin, K.: Adaptive structuring of distributed DBs. In: Proceedings of the National Computer Conference (AFIPS 1982), June 7-10, pp. 691–696. ACM, New York (1982)
30. PostgreSQL 8.4.18 Documentation, ch. 26. Monitoring DB Activity (2013), http://www.postgresql.org/docs/8.4/static/monitoring-stats.html
31. Hadri, B., Fahey, M., Jones, N.: Identifying software usage at HPC centers with the automatic library tracking DB. In: Proceedings of the 2010 TeraGrid Conference (TG 2010), Article 8, 8 pages. ACM, New York (2010)
32. Novais, R.L., Torres, A., Mendes, T.S.: Software evolution visualization: A systematic mapping study. Information and Software Technology 55, 1860–1883 (2013)
33. German, D.M., Hindle, A., Jordan, N.: Visualizing the evolution of software using SoftChange. In: Proceedings SEKE 2004 The 16th International Conference on Software Engineering and Knowledge Engineering, pp. 336–341 (2004)
34. Storey, M.A.D., Cubranic, D., German, D.M.: On the Use of Visualization to Support Awareness of Human Activities in Software Development: A Survey and a Framework. In: Proceedings of the 2005 ACM Symposium on Software Visualization, pp. 193–202. ACM (2005)
35. Mercer, E., Person, S., Rungta, N.: Computing and Visualizing the Impact of Change with Java PathFinder Extensions. SIGSOFT Softw. Eng. Notes 37(6), 1–5 (2012)
36. Störrle, H.: Making Sense of UML Class Model Changes by Textual Difference Presentation. In: Proceedings of the 6th International Workshop on Models and Evolution, ME 2012, New York, NY, USA, pp. 3–8 (2012)
37. Xing, Z., Stroulia, E.: UMLDiff: An Algorithm for Object-Oriented Design Differencing. In: Proceedings of the 20th IEEE/ACM International Conference on Automated Software Engineering, ASE 2005, pp. 54–65. ACM, New York (2005)
38. OMG: Documents Associated With Unified Modeling Language (UML), V2.4.1, http://www.omg.org/spec/UML/2.4.1/
39. Krogstie, J.: Model-Based Development and Evolution of Information Systems. Springer, London (2012)
40. Frankel, D.S.: Model Driven Architecture: Applying MDA to Enterprise Computing. Wiley Publishing, Indianapolis (2003)
41. Visual Paradigm. Compare Logical and Physical ERD, http://www.visual-paradigm.com/highlight/compare_logical_physical_erd.jsp
42. Grose, T.J., Doney, G.C., Brodsky, S.A.: Mastering XMI Java Programming with XMI, XML, and UML. John Wiley & Sons (2002)

43. Sen, A.: Metadata management: past, present and future. Decision Support Systems 37(1), 151–173 (2004)
44. I.A.: technical discussion of software modeling, http://public.dhe.ibm.com/software/rational/web/whitepapers/ValueOfModeling.pdf
45. Aguilar, R., Pan, J., Gries, C., San Gil, I., Palanisamy, G.: A flexible online metadata editing and management system. Ecological Informatics 5(1), 26–31 (2010)
46. Wootton, C.: Extracting Metadata from Word Documents. Focal Press (2007)
47. W3Schools. XML Schema Tutorial, http://www.w3schools.com/schema/
48. Castro-García, L., López-Morteo, G.: An international analysis of the extensions to the IEEE LOMv1. Computer Standards & Interfaces 35(1), 567–581 (2013)
49. Lu, E.J.-L., Wu, B.-C., Chuang, P.-Y.: An empirical study of XML data management in business information systems. Journal of Systems and Software 79(7), 984–1000 (2006)
50. Goswami, S., Kundu, C.: XML based advanced distributed DB: implemented on library system. International Journal of Information Management 33(1), 28–31 (2013)
51. Lee, K.K.-Y., Tang, W.-C., Choi, K.-S.: Alternatives to relational DB: Comparison of NoSQL and XML approaches for clinical data storage. Computer Methods and Programs in Biomedicine 110(1), 99–109 (2013)
52. Faisal, S., Sarwar, M.: Temporal and multi-versioned XML documents: A survey. Information Processing & Management 50(1), 113–131 (2014)
53. Haw, S.-C., Lee, C.-S.: Data storage practices and query processing in XML DBs: A survey. Knowledge-Based Systems 24(8), 1317–1340 (2011)
54. Object Management Group: OMG MOF 2 XMI Mapping Specification. Version 2.4.1 (2013), http://www.omg.org/spec/XMI/2.4.1
55. Oracle. The J2E 1.4 Tutorial. Transforming XML Data with XSLT, http://docs.oracle.com/javaee/1.4/tutorial/doc/JAXPXSLT6.html
56. Objects By Design. Transforming XMI to HTML, http://www.objectsbydesign.com/projects/xmi_to_html.html
57. Muench, S.: Chapter 7. Transforming XML with XSLT. In: Building Oracle XML Applications, pp. 275–309. O'Reilly Media (2000)
58. Selmeci, A., Orosz, I., Orosz, T., Gyorok, G.: ERP change management for innovation and sustainability applied to User Interfaces. In: IEEE 16th International Conference on Intelligent Engineering Systems (INES), June 13-15, pp. 341–346. IEEE (2012)
59. Ng, C.S.P., Gable, G., Chan, T.: An ERP maintenance model. In: Proceedings of the 36th Annual Hawaii International Conference on System Sciences, January 6-9, p. 10. IEEE (2003)
60. Jamshidi, A., Rahimi, S.A., Bartolome, A.R., Ait-kadi, D.: A new framework for risk assessment in ERP maintenance. In: Annual Reliability and Maintainability Symposium (RAMS), January 27-30, pp. 1–6. IEEE (2014)
61. Kwahk, K.-Y.: ERP Acceptance: Organizational Change Perspective. In: HICSS 2006. Proceedings of the 39th Annual Hawaii International Conference on System Sciences, January 04-07, vol. 8, p. 172. IEEE (2006)
62. Rajapakse, J.: Can ERP Adoptions Change Organisational Culture in Developing Countries in Asia? An Empirical Investigation. In: 45th Hawaii International Conference on System Science (HICSS), January 4-7, pp. 5093–5101. IEEE (2012)

Evaluating the Application of Interactive Classification System in University Study Course Comparison

Ilze Birzniece[1], Peteris Rudzajs[2], and Diana Kalibatiene[3]

[1] Study Department, Riga Technical University, Latvia
[2] Institute of Applied Computer Systems, Riga Technical University, Latvia
[3] Department of Information Systems, Vilnius Gediminas Technical University, Lithuania
{ilze.birzniece,peteris.rudzajs}@rtu.lv,
diana.kalibatiene@vgtu.lt

Abstract. Large amount of routine work necessary to perform comparative analysis of university courses possess the importance of the automation of this task. Possibility of effortless detection of similarity between different courses would give the opportunity to organize student exchange programmes effectively and facilitate curriculum management and development. The application of smartly adapted machine learning technologies in long term could reduce the manual course comparison effort. The goal of this paper is to present the application of earlier proposed inductive learning based classification system (accompanied with interactive capabilities) to directly and indirectly compare study courses semi-automatically. The evaluation of the proposed system has been carried out in 4 consecutive experiments which proved the ability to decrease the number of misclassified instances if uncertain classifications are detected and passed to the expert's review.

Keywords: Machine learning, interactive classification, inductive learning, curricula comparison.

1 Introduction

Nowadays there are plenty of different study courses provided by different higher education institutions around the world. Here the problem of detecting similarities between them arises. The comparison of study programmes and courses is necessary in several educational tasks. One of them is student mobility. Taking into consideration the number of different education providers this is a time consuming task if maintained only by human expert. Although one of the main features of the Bologna process is to encourage creation of a common model for Higher Education in Europe [1], there still does not exist a generally established standard for describing study courses in all European universities, and they currently appear both as semi-structured and unstructured textual descriptions. This fact creates the main difficulty for course comparison automatically. Therefore, in reality comparison of study programmes and individual courses is a task that is performed manually, although machine learning techniques could be of great value in reducing human routine

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 335–346, 2014.

activities. Machine learning algorithms learns classification from training examples (past experience) and uses induced classifier for dealing with new instances of the same problem area.

In this paper the authors present the application of their earlier proposed interactive classification system for study course comparison. The rest of the paper is structured as follows. Section 2 presents related work. Section 3 briefly describes the framework for developing Interactive Classification System (InClaS). Section 4 describes detailed experiment on applying InClaS for university course comparison and Section 5 concludes the paper.

2 Related Work

The necessity to compare educational documents[1] appears in different forms and can be conditionally divided into three categories ([2–7]): (1) student exchange programmes, (2) new curriculum development and (3) teaching material and learning object categorization for, e.g. e-learning systems. In the scope of this paper we consider only the first category, namely, mutual comparison of course content.

Study course description most often is a semi-structured text which usually includes sections like “prior knowledge”, “learning outcomes”, etc. It is important to distinguish between these sections. Besides, a semi-structured text has a significantly richer and more complicated structure than a plain-text, and the relation among semi-structured documents is harder to be fully utilized if only text categorization is used [8, 9].

Existing research in the area of curricula comparison does not solve the problem of study course comparison. It has been proposed to represent study programmes as concept maps [3] and a system based on schema matching [10]. In this approach curricula are compared according to their structure. However, one of the basic tasks in comparing curricula is the comparison of individual courses in the course content level that has not been included in this research. Authors of [2] present an unsupervised classification mechanism, which organizes educational documents from e-learning system into clusters. Authors of [6] describe design of methodology for classification of learning objects which can appear in different forms, e.g. course outlines, transcripts, etc., without well-defined metadata. Classification of a new learning object is done by finding the smallest distance to the cluster, where clusters define subdomains of interest. However, in [2] and [6] it is not the course description that is used as the input. Both of these approaches also assume that objects relate to only one category, although in practice it is not the case when comparing documents in distinct curricula. Comparative analysis of educational documents is a complicated task both for experts and computer systems. Therefore automation of this process requires specific approaches and expert participation. Semi-structured document representation requires the use of various information extraction methods. Authors of Academic e-Advising system [5] point out that system’s results could undoubtedly be

[1] A term *educational document* is used to denote different types of materials for educational content and assessment, including course descriptions, teaching materials, academic credentials, etc.

improved by expanding the size of the training corpora and involving an expert. The system would also benefit from the implementation of an easy mechanism for manual inspection and augmentation of the extracted data to improve data quality for further use. Therefore, course comparison task proves to be non-trivial for application of machine methods directly because of mixture of domain features. The problem domain can be defined by the following properties (expanded in [11]).

- Understanding decision making steps is important for expert to trust the results produced by the computer system
- Small initial experience/learning base (1-10 examples for each class with around 38 attributes)
- Many (10-50) classes with similar probability to appear (number of courses within one study programme)
- Multi-label class membership (a certain course can be mapped to several other courses)
- Semi-structured data (course descriptions)
- Classification results are needed for a system's user, not only for a system, therefore, need to be interpretable

To satisfy these needs and propose semi-automatic approach for supporting study course comparison, an interactive inductive learning based classification system is developed and its experimental evaluation is presented in this paper.

3 Framework of Interactive Inductive Learning Based Classification System (InClaS)

According to the related work and our researches the framework for developing interactive inductive learning based classification system (InClaS) was proposed. Since the aim of this paper is to present experiment results of InClaS, we present only short description of the framework. For more details see publications [12–15]. InClaS framework defines algorithms, methods, approaches and architectures for developing semi-automatic classification system which allows interactivity with an expert at the classifier's applying stage if the classifier meets an instance which it cannot classify or is not confident of the classification made. Usage of such a system is *feasible* in areas where human-expert is available. The interactive classification approach is *more appropriate* than the automatic classification in areas where (1) it is essential to receive a correct classification for as much instances as possible, (2) it is hard to extract or define domain features resulting in attributes which do not describe the underlying concept completely and/or (3) only a small initial learning set is available and it is suspected not being representable.

The framework of InClaS consists of three levels (see Fig. 1). They are as follows:

1. Generic model
2. Model for multi-label classification
3. Prototype

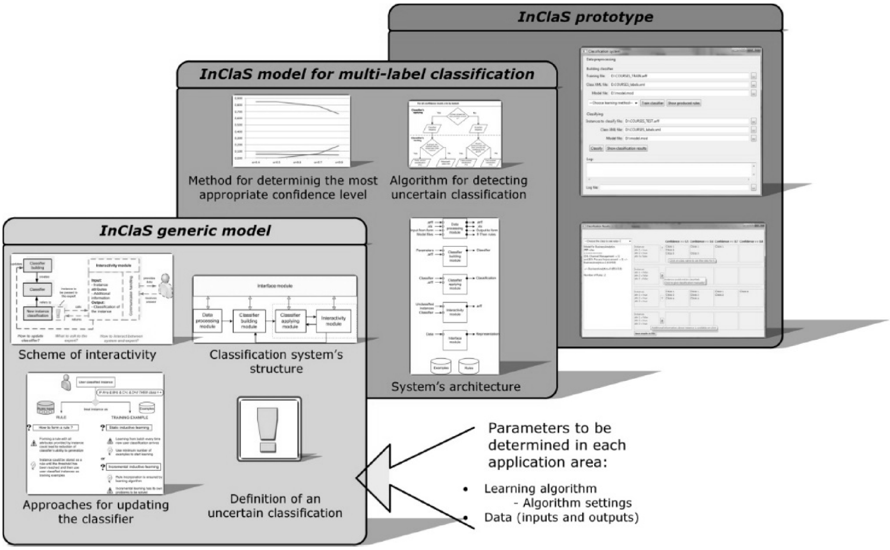


Fig. 1. Framework of InClaS (originally published in [15])

Generic model presents an interactivity model which ensures expert involvement if the classifier meets an object that cannot be classified or is not confident to the classification made, various components extending traditional classification system are developed (see more in [12–16]).

To deal with multi-label classification, InClaS model has been extended with the following components: *Algorithm for detecting uncertain classification*, *Method for determining the most appropriate confidence level* for each data set and *Architecture of a classification system.*

The developed InClaS prototype is based on the components described in the first two levels of the framework, as well as 11 basic static learning algorithms from Weka software [17] and multi-label classification methods implemented in Mulan [18] library. In InClaS prototype data input and output is provided through graphical user interface (GUI). The classification system extracts and saves the rules held in the classifier (in a text file) in a human-readable form. To implement an interactivity scheme, the classifier's application stage has been improved with the ability to trace the confidence of classification and intercept uncertain classifications. Classification results are presented to the user (expert), which can apprise classes assigned with different confidences and make his classification if no classification is given with the confidence 0.5 or more. Thus all together the InClaS prototype provides a unique environment for multi-label classification in a more user-friendly way than it was possible before as well as novel interactivity facilities between the classification system and its user.

Novelty of the proposed approach lies in selecting appropriate methods known in the area of machine learning and featuring inductive learning with interactivity to create a machine learning based solution for a domain where computer-based techniques so far do not prevail.

4 InClaS for University Study Course Comparison

If we want to define the study course comparison as a classification task, we need to define attributes and classes. The task of machine learning algorithm in this case is to adopt experience from a human expert in deciding whether two courses are similar. There are two main stages of classifier formation and application for study course comparison. Initially there are courses that we want other courses to compare to; they are called *target courses* and form classes. Courses are described with attributes extracted from their description (approaches for attribute extraction will be described in more details in next section). For target courses expert's classification is the same as the course title.

UML activity diagrams are used to describe the general steps in (1) preparing interactive classification system with providing expert-classified study course descriptions and (2) applying induced rule-based classifier for detecting similarities between previously unseen study courses.

To train the classifier, a human expert should classify also some courses from other curricula. The expert can decide that the unknown study course refers to one or more than one target courses. The system receives course attributes and expert's classification and tries to find correspondences, and infer rules from the provided examples (see Fig. 2.).

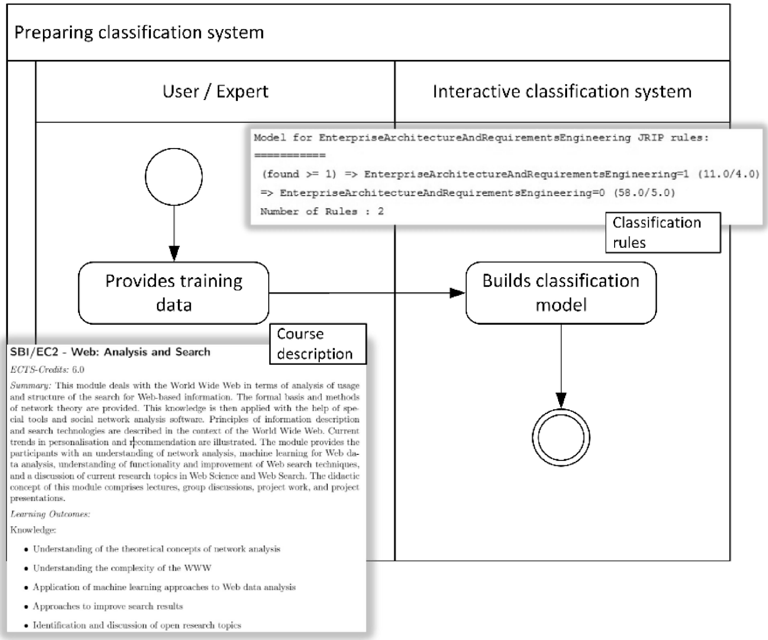


Fig. 2. The process of training classifier

After the training phase the classifier can be put into action. If the classifier is able to find an appropriate rule in its rule base, a decision of course class (or classes) is made. If no rule can classify the course, an expert is asked to classify it. The expert decision can be further used to improve the rule base since human-given advice can be saved and formed as a new rule to be incorporated into an existing rule base. This approach allows starting use of a classifier which is not fully trained, e.g. because of lack of expert labelled training examples (see Fig. 3).

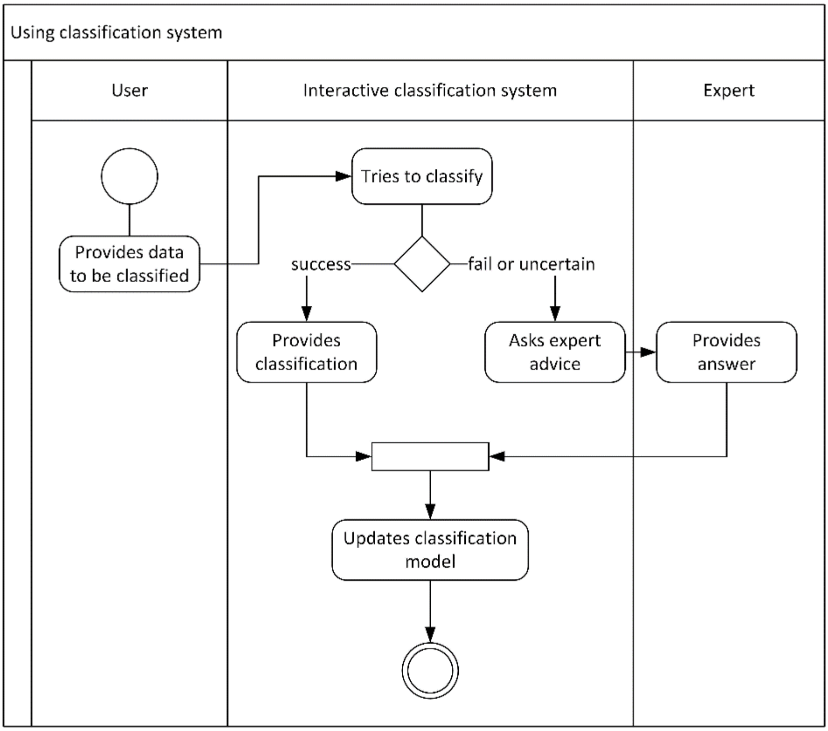


Fig. 3. The process of using classifier

5 Evaluation of InClaS in Education

This section describes the experiment and main results in practical application of InClaS in the domain of higher education. The aim of experiments is to examine the utility of the InClaS framework, usability of the system's prototype and evaluate the impact of chosen settings to study course comparison task.

In order to assess an InClaS utility the number of misclassified instances, applying the standard non-interactive approach and the proposed interactive approach is compared. Regarding usefulness of the proposed solution in education area the following aspects are evaluated:

- Verification that this problem domain is not appropriate for traditional automatic machine learning solutions, whereas inductive learning methods based interactive multi-label classification system for supporting study course comparison can provide acceptable solution.
- Evaluation of a direct (using attributes achieved directly from full course descriptions) and indirect (using mediated attributes from course descriptions) study course comparison.

5.1 Data and Experimental Settings

In this experiment the full data set consists of 79 instances (study course descriptions) from different European universities providing Business Informatics related curricula, namely, 25 instances from Riga Technical University, which serve as target courses or classes, whereas 54 other courses are mapped to target courses (6 instances from University of Rostock, 31 from Vienna University of Technology and 17 from University of Vienna). In a reduced set, the classes which are represented with less than 4 instances are removed. For practical experiments in university course comparison two attribute extraction approaches are chosen. For direct course comparison, text classification approach is applied which makes use of word vectors obtained from full course descriptions. Indirect course comparison involves mediating framework (European e-Competence Framework [19]) for extracting semantically meaningful information from course descriptions – learning outcomes defined as competences, number of credit points, study level. For more detailed description of attribute selection and preprocessing, see previous publications [13, 15].

Parameters of data sets are given in Table 1.

Table 1. Study course data set

	No. of attributes	**No. of instances**	**No. of classes**
Full data set (based on word vectors)	1884	131 (79)	25
Full data set (based on competencies)	38	79	25
Reduced data set (based on competencies)	38	64	12

Experimental settings are described in Table 2. Four setting combinations or distinct stages of experiments are defined: (1) word vectors with automatic classification, (2) mediated attributes with automatic classification, (3) word vectors with InClaS, and (4) mediated attributes with InClaS.

Experiments include a wide range of multi-label classification methods, mostly different algorithms and meta-algorithms that go beyond initial problem transformation as binary relevance (see [13] for more information). The evaluation of methods that learn from multi-label data requires different measures than those used for single-label data [20]. To evaluate classification performance, five popular measurement functions are chosen. Hamming loss is an example-based measure which computes the percentage of labels whose relevance is predicted incorrectly. It shows the average binary classification error. Micro-averaged precision and recall are multi-label alternatives for corresponding binary evaluation measures. To evaluate ranking, two measures are chosen. One-error says how many times the top-ranked label is not in the set of relevant labels of the instance, while the coverage calculates the average of how far we need to go down the ranked list of labels in order to cover all the relevant labels of the example.

To consider usefulness of user involvement in classification process and impact to number of misclassified instances, InClaS framework introduces several simple measures to be detected and evaluated:

Partly correct or completely correctly classified instance (PC) – at least one of predicted classes is the actual class of an instance, $Y_i \cap Z_i \neq \emptyset$, where

Y_i – actual label set of instance i, Z_i – predicted labels set of instance i.

Misclassified instance (M) – none of predicted classes is the actual class of an instance, $Y_i \cap Z_i = \emptyset$.

True uncertain classification (TU) – the classifier would misclassify an instance (*M*) (that is, with the confidence level 0.5 none of actual classes would be predicted).

False uncertain classification (FU) – the classifier would classify instance partly or completely correctly (*PC*) (that is, with the confidence level 0.5 at least one of actual classes would be predicted).

Table 2. Experimental settings for study course comparison

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Input data set	Full study course descriptions (extracting word vectors in preprocessing)	Competencies of study course, number of credit points, study level	Full study course descriptions (extracting word vectors in preprocessing)	Competencies of study course, number of credit points, study level
Classification approach	Automatic classification	Automatic classification	Interactive classification (InClaS model)	Interactive classification (InClaS model)
Classification algorithms (methods)	20 classification algorithm-method combinations (from *Weka* and *Mulan*)	20 classification algorithm-method combinations (from *Weka* and *Mulan*)	4 best methods from Experiment 1	4 best methods from Experiment 2
Evaluation measures	Hamming loss, Micro-average precision, Micro-average recall, One-error, Coverage	Hamming loss, Micro-average precision, Micro-average recall, One-error, Coverage	*M, PC, FU, TU*	*M, PC, FU, TU*

5.2 Main Experimental Results

Experiments 1 and 2 evaluates twenty different classification methods (methods MC-Copy, MC-Ignore, IncludeLabels, LP, MLStacking) or method combinations (Binary relevance (BR) with Naïve Bayes, KStar, IBk, Bagging, Stacking, AdaBoost, Part, Prism, JRip, REPTree, RF, J48; RAkEL with J48) on the full study course data sets without applying interactivity by means of Hamming loss, Micro-average precision, Micro-average recall, One-error, Coverage. Four method combinations which received four best evaluations in 10-fold cross-validation during experiment 1 and 2 are further used in experiments 3 and 4.

Experiment 3 evaluates four methods which achieved the best results in Experiment 1 in 3 times repeated random sub-sampling validation on word-vector based data set. Results in Table 3 should be interpreted as follows. Using the automatic classification, where only partly or completely correct classifications and misclassifications exist, 27% of instances would be partly correct (PC) (in case of RAkEL method) and 73% – misclassified. If the interactive approach is used, the number of PC remains the same; however, 33% of instances from previously misclassified are marked as uncertain to the classifier and given to the expert, reducing the number of misclassified instances to 40%. Without applying interactivity the number of misclassified instances is much higher for all methods. Note the assumption that the expert makes correct classifications to the instances passed to him. To all appearances, the given data set does not provide a complete concept description as it was assumed when considering domain features.

Table 3. Interactive approach for direct study course comparison (word vectors)

	Partly correct (*PC*)	**True uncertain classification (*TU*)**	**False uncertain classification (*FU*)**	**Misclassified (with interactivity)**	**Misclassified (without interactivity)**
RAkEL(J48)	0.267	0.333	0.000	0.400	0.733
BR(AdaBoost)	0.100	0.400	0.000	0.500	0.900
BR(Bagging)	0.067	0.600	0.000	0.333	0.933
BR(JRip)	0.267	0.367	0.000	0.366	0,733

Table 4 represents results of Experiment 4, where four best method combinations from Experiment 2 are applied to competence-based study course data set.

Table 4. Interactive approach for indirect study course comparison (competencies)

	Partly correct (*PC*)	**True uncertain classification (*TU*)**	**False uncertain classification (*FU*)**	**Misclassified (with interactivity)**	**Misclassified (without interactivity)**
BR(NB)	0.234	0.633	0.000	0.133	0.766
BR(Bagging)	0.167	0.733	0.000	0.100	0.833
BR(AdaBoost)	0.267	0.433	0.000	0.300	0.733
BR(JRip)	0.267	0.367	0.000	0.366	0,733

Alike the results of Experiment 4, the ability of the InClaS to track uncertain classifications allows to decrease the number of misclassified instances, although results vary much between the methods used. Graphical representation of JRip results in Fig. 4 emphasizes the impact of the interactive approach even more. Without interactivity (part A of Fig. 4), 73% of instances would be misclassified reaching only 27% of PC. Such classification results do not encourage the use of the automatic classification in this problem domain. In turn, the interactive approach (part B of Fig. 4) with the ability to handle uncertain classification makes it possible to save half of misclassified instances and assign to them correct classifications after the expert's review. Thus, 37% of instances are misclassified, which, obviously, is not a great result, but is much more promising than 73% with the automatic classification.

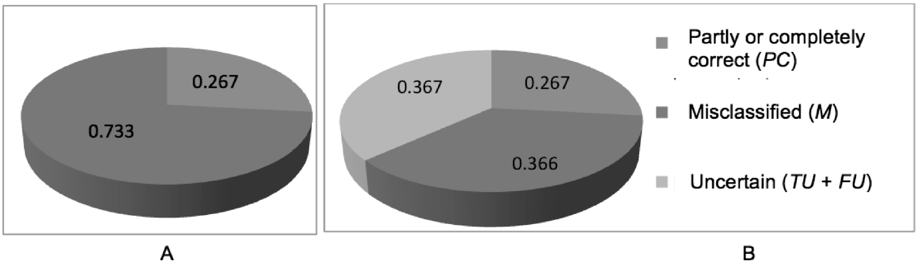


Fig. 4. Test results of JRip algorithm with automatic (A) and interactive (B) classification

To consider the situation when the number of training examples regarding each class has increased, experiments with the reduced data set are carried out. The results lead to conclusion that interactive classification system improves its results and less frequently disturbs the expert when the training set grows in time. Therefore, it is useful to spend expert's time more in the initial period of classifier's usage in order to obtain better classification results later. Figure 5 shows the difference between results in the data set with reduced number of classes where each class is described with slightly higher number of examples (part A) and the full data set, which includes many underrepresented classes (part B). In reduced data set PC reach 50% of instances leaving 17% of instances for expert's decision and also decreasing the number of misclassified instances. All these parameters are improved in comparison to the initial data set.

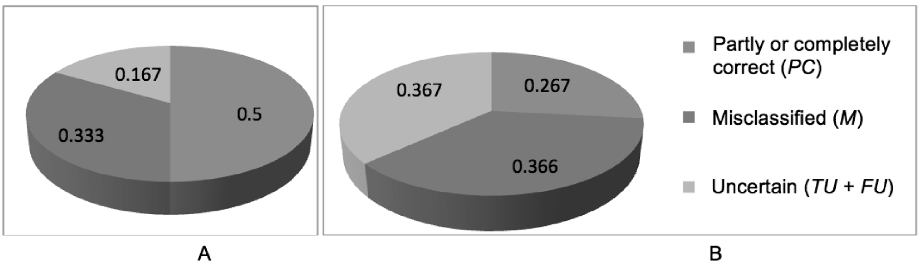


Fig. 5. Test results of JRip algorithm task with reduced (A) and full (B) course data set

Experimental results also deny assumption that the indirect (competence-based) course comparison provides better classification results than the direct (word vector based) comparison. That is, structured and meaningful information extraction from course descriptions produce attributes, which do not surpass full course description usage to make word vector based attributes by means of number of misclassified instances and (partly) correct classifications. Both approaches can be used, however, the indirect comparison currently requires much more expert's work in attribute extraction phase since competencies are not accessible directly in course descriptions. If course descriptions are standardized, it makes the situation more convenient for such approach. As a disadvantage of word vector usage to define attributes its low semantic meaning should be mentioned. It does not provide useful knowledge to the expert as it only describes occurrences of different words in descriptions wherever in the text they appear – either preconditions or learning outcomes. Therefore, the knowledge about underlying communalities of the course content can be mined if meaningful attributes are used, like competencies which the study course provides.

6 Conclusions and Future Work

InClaS framework defines algorithms, methods and other components which allow to develop an interactive classification system for decreasing the number of misclassified instances in domains where a human-expert is available. In this paper we evaluated interactive multi-label classification system's prototype developed in accordance to this framework in the domain of education. This domain is inappropriate for traditional automatic machine learning applications due to its complex nature – lack of well-defined comparison set, hard-to extract attributes, potential one-to-many course correspondences. In order to overcome these difficulties and reduce expert workload as well, a semi-automatic or interactive approach is proposed.

Course correspondences between *Business Informatics* master study programme in *Riga Technical University* and courses of several corresponding study programmes in Europe are detected to constitute data set. Evaluation of the InClaS has been carried out in 4 consecutive experiments which proved the ability to decrease the number of misclassified instances if uncertain classifications are detected and passed to the expert's review. Improvements of interactive approach are especially significant in the areas like study course comparison where the automatic classifier is weak and without applying the interactive approach it would not be possible to apply machine learning based classifier at all due to the many incorrect decisions it produces. Therefore, to reduce manual expert involvement in the study course comparison semi-automatic classification system could form a core engine in a wider decision-support system.

Acknowledgements. This work has been supported by the European Social Fund project Nr. 2013/0012/1DP/1.1.1.2.0/13/APIA/VIAA/051.

References

1. Kennedy, D., Hyland, Á., Ryan, N.: Writing and Using Learning Outcomes: a Practical Guide. In: EUA Bologna Handbook: Making Bologna Work (2009)
2. Alves, H., Figueira, Á.: A Educational Library based on Clusters of Semantic Proximity. In: IADIS European Conference on Data Mining, pp. 226–228 (2011)
3. Anohina-Naumeca, A., Graudiņa, V., Grundspeņķis, J.: Curricula Comparison Using Concept Maps and Ontologies (2012)
4. Birzniece, I., Kirikova, M.: Interactive Inductive Learning Service for Indirect Analysis of Study Subject Compatibility. In: Proc. BeneLearn 2010, Belgium, Leuven, pp. 1–6 (2010)
5. Biletskiy, Y., Brown, J.A., Ranganathan, G.: Information extraction from syllabi for academic e-Advising. Expert Syst. Appl. 36, 4508–4516 (2009)
6. Ranganthan, G.R., Biletskiy, Y., MacIsaac, D.: Machine Learning for Classifying Learning Objects. In: IEEE CCECE/CCGEI, pp. 280–283 (2006)
7. Biletska, O., Biletskiy, Y., Li, H., Vovk, R.: A semantic approach to expert system for e-Assessment of credentials and competencies. Expert Syst. Appl. 37, 7003–7014 (2010)
8. Sebastiani, F.: Machine learning in automated text categorization. ACM Comput. Surv. 34, 1–47 (2002)
9. Yang, J., Chen, X.: A semi-structured document model for text mining. J. Comput. Sci. Technol. 17, 603–610 (2002)
10. Saleem, K., Bellahsene, Z., Hunt, E.: PORSCHE: Performance ORiented SCHEma mediation. Inf. Syst. 33, 637–657 (2008)
11. Birzniece, I.: Interactive Use of Inductive Approach for Analyzing and Developing Conceptual Structures (2012)
12. Birzniece, I., Rudzajs, P.: Machine Learning Based Study Course Comparison. In: IADIS Conference on Intelligent Systems and Agents (ISA 2011), Italy, Rome, pp. 107–111 (2011).
13. Birzniece, I.: Machine Learning Approach for Study Course Comparison. International Conference on Machine Learning and Data Mining (MLDM 2012), pp. 1–13. IBai, Berlin (2012)
14. Birzniece, I.: Interactive Inductive Learning System. In: Barzdins, J., Kirikova, M. (eds.) Selected Papers from the DB&IS 2010, pp. 380–393. IOS Press, Latvia (2010)
15. Birzniece, I.: Development of Interactive Inductive Learning Based Classification System's Model (2013)
16. Birzniece, I., Kirikova, M.: Interactive Inductive Learning: Application in Domain of Education. Sci. J. Riga Tech. Univ. Comput. Sci. Appl. Comput. Syst., 1–8 (2011) (in press)
17. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA Data Mining Software: An Update. SIGKDD Explor 11, 10–18 (2009)
18. Tsoumakas, G., Spyromitros-Xioufis, E., Vilcek, J., Vlahavas, I.: Mulan: A Java Library for Multi-Label Learning. J. Mach. Learn. Res. 12, 2411–2414 (2011)
19. European Committee For Standardization: European e-Competence Framework 2.0, http://ecompetences.eu/
20. Tsoumakas, G., Katakis, I., Vlahavas, I.: Mining Multi-label Data. In: Maimon, O., Rokach, L. (eds.) Data Mining and Knowledge Discovery Handbook, pp. 667–685. Springer, US (2010)

Information Standards Enabling or Constraining Innovative Hospital Facilities? -A Scandinavian Case

Christian Koch[1], Kim Jacobsen[2], and Anita Moum[3]

[1] Chalmers University of Technology, Gothenburg, Sweden
[2] K-Jakobsen A/S, Copenhagen, Denmark
[3] Norwegian Institute of Science and Technology, Trondheim, Norway

Abstract. The losses from suboptimal interoperability in IT supported design, production, and operation of health care facilities are tantamount. In these years new built of hospitals in Scandinavia could be realized in a more efficient and innovative way if using information standards. This paper inquires into whether computerized information standards enable or constrain innovation in public procurement of buildings. In architectural and engineering design of public buildings the project based product development tends be done in constellations of firms in interorganisational contracting, which do not provide stability or room for innovation. A large hospital project was investigated through interviews, documents and observations. The effects of implementing building information standards are both inter- and intraorganisational. The building client claims to have saved money, through better structured building component data that gave considerable positive effects during tendering. The IT-suppliers develop IT-tools, and the AEC companies can commence develop services preparing for new markets.

Keywords: information standards, innovation, hospital buildings, Denmark, Norway, Sweden.

1 Introduction

The healthcare sector represents an important part of the Scandinavian welfare states [1] and these countries invest the most in Europe [2]. And in these years we witness an unprecedented wave of investment in renewal and extension of this infrastructure. Norway thus invested an estimated 10 billion Euro in new hospitals from 2000 to 2011 [3], and Denmark has announced future investment at 5, 5 billion Euro[4]. Sweden as well invests in large projects in its major cities and generally [2].

A part of this development is the emergence of large business suppliers of services for building and operating these hospital services in architectural, engineering, contracting and real estate industries [5]. Companies like Arkitema (architects), Bravida (technical installations), C.F. Møller (architects), Cowi (engineers), Skanska (contractors) and Veidekke (contractor) all contribute to the renewal of the Scandinavian hospitals and do it in a multinational manner across the Scandinavian

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 347–361, 2014.

countries and beyond. These companies appear well placed to participate in further development of healthcare sectors in for example US, and BRIC-countries.

However in order to exploit these business opportunities Scandinavian architects, consulting engineers, contractors (AEC) and real estate firms need to tackle the issue of lack of proper interoperability of building information, which continues to be a major challenge not only for hospital projects but for the building industry as such [6, 7]. As the information get incorporated in Building Information Models (BIM) and when these complex combined 3D geometric, and multidimensional datamodels become used in design, production and operations processes of the large complex hospital buildings as central design- and support tool, a swift, efficient and smoothless transfer between different parts of the process is crucial [8, 9]. Moreover an standardized internal structure of data in building information models not only enable transfer/interoperability but also help using BIM in a strategy for creating innovative solutions for new sustainable energy efficient buildings and to enable competitive innovations for Scandinavian building companies, enabling their prospering both locally and globally.

There is thus a need for common tools, standards and work methods. Design production and use of state-of-the-art sustainable building norms similarly requires a set of calculation, design and monitoring tools, which today remains national in their character in contrast to the demands of the market. At present international standards such as Eurocodes and Industry Foundation Classes (IFC) [10] and defacto standards such as DWG only partially cover the Scandinavian building processes, even if IFC is gradually developing as a global standard.

One nationally embedded, but new, standardisation tool is the Cuneco Classification System (CCS) [11]. This clustered set of building information standards is developed in a Danish context, but designed to be used in international building sectors. The standardisation of CCS target building information as used in design, building and operation of buildings. CCS provide a classification of rooms as basic elements of the building product, automated identification and classification of building components, building sub systems, properties of building component, classification of resources for the building process and a standard for information level supporting the building process. A central element is a cloudbased server for access to the standards and the standards are currently built into a small group of software vendor's product. The computerized CCS enables smoothless data transfer between software packages which today are often not interoperable, be it software for engineering design of heating- ventilation- and air condition (HVAC) and energy calculation- software, when combining sustainable design and indoor climate design. Also CCS would, through more integrated IT infrastructures, enable virtual collaboration and coordination which gain importance in and outside the Scandinavian region. CCS would thereby systematically overcome obstacles for innovation in building design, production and operation. Yet and importantly it has still not been tested outside Denmark and it is more than likely that its international diffusion will be complete dependable of design of engines and interfaces that can translate from one standard to another in a future multi standard landscape, such as it is known in other areas of standardisations [12,13].

The aim of this contribution is to investigate whether computerized information standards using CCS as case enable or constrain innovation of hospital buildings and investigate under what circumstances information standards would have an improved impact on innovation.

The two main elements of this investigation are a preliminary literature review and a case of a Danish hospital testing the concrete set of standards, i. e. CCS [14].The standards are presently released in their first operable version. Other parts of the backup structures are still under final development. The classification is expected be fully avaible for use by the end of 2014 [11].

In the context of hospital building, the public authorities have often been change agents using procurement to generate standardisation and innovation. Recently EU has pointed to the public authorities as spearhead for energy efficient buildings mitigating climate change [15]. Public procurement of hospital buildings involves demands of use of IT, collaboration and quality tools. This context is therefore particularly interesting for studying links between standards and innovation.

A literature review is used to develop five types of impacts that standards can have on innovation. This is thus the papers primary contribution to analyse how the envisaged standardisation with CCS enables innovation in procurement of hospitals and in building in general. The five types of impact is investigated in the case and the effects presented. However also the barriers for obtaining these results of improved innovation are identified. The paper is structured in the following way. After a method presentation, a selective literature review is used to develop the five types of innovation enhancing effects of standardisation. Using this framework structures the presentation of the Danish Hospital case, which is analyzed in the discussion arriving at the conclusion.

2 Method

The scientific approach taken here for business informatics research is critical interpretivist [16, 17] adopting a sociomaterial view on innovations and standards [18]. This approach emphasizes the inseparability of the social and technical aspect as well as both aspects negotiated interpreted character. The results presented here are preliminary as the research project behind is ongoing. The method consists of two main elements, a literature review and a case study, followed by a qualitative analysis:

The *literature review* is carried out covering Science Technology and Society studies, innovation studies, information systems, Human Computer Interaction and organisation theory approaches [17, 19]. To do this several search engines was used, including ABI/Inform and Scopus. Main search item was “standards AND innovation”. Through several rounds of iteration this gave a focus set of roughly 50 articles. Also backtracking and an element of serependity were involved [20] when combining searches in several communities. Currently 55 articles and books are under review, and 31 have been used for this paper, a selection based on the preliminary reading of the abstracts of the articles and introduction of the books. The review leads to a framework of understanding of the relation between computer based standardized information and five types of innovation.

The selected domain for the *case study* of effects of standards on innovation is architectural and engineering design, production and operation of hospital buildings. Here handling of information gives rise to substantial interoperability issues that tends to hamper innovation. Moreover the project based production of products in the building sector tends be done in constellations of firms in inter organisational contracting, which do not provide stability or room for innovation. The Architecture, Engineering and Construction industry is therefore often portrayed as lagging behind regarding innovation. Here a hospital project with its size in resources and time potentially provide an exemption that could foster innovation. As empirical "test" field a large Danish hospital project is used, which is the building project in Denmark which have used the CCS standards most extensively. It is therefore a critical case and not intended to be typical. The material gathered for this case is composite. First the second author took part in a test project, where building consultants, IT vendors and the client participated and the author could carry out observations through presence at 40 meetings with the design team, the steering group, the IT-supplier's groups and other events. He also carried out a series of interviews, typically of the duration of one hour with these actors in the test project. Second, and supplementary, the first author acts as process evaluator for "Knowledge Center for increased Productivity and Digitalization in Construction" (Cuneco) [11], together with two colleagues from 2010-2015. The process evaluation of the center encompasses by February 2014, seven short half-year process evaluation reports. As the entire center focuses on establishing an building information standardization, most of the material gathered is relevant as background material to the research questions raised here. Data collection encompasses interviews (42), participant observation of events (17), document analysis (141). However more specifically a representative of the hospital test project was interviewed four times during 2012 and 2013. And the document analysis of the developed standards contributed to the understanding of the innovation opportunities. The third author has an independent position vis a vis the Danish classification effort.

The case material is *analysed* with the framework developed by the literature review. Some of the dimensions turn out to have more weight than others, and it is therefore not strived at to give the dimensions equal emphasis. The analysis is moreover qualitative even if a few figures do occur. Finally the analysis carried out here builds on [21, 22].

The limitations of the study are that the case of classification studied is not a long term stabilised one, but was rather prototypes under development. Many of the more indirect innovation types relating to business models and community [23] is more of a future potential for the time being. Moreover two authors are deeply involved in the development of CCS and one author has been project management for the hospital test project central is the paper. The closeness is however seen at a time as a strength and a weakness as it provides detailed insight in the processes of the case. The third author has acted as critical external vis a vis these indepth insights.

3 Review of the Literature

This selective literature review follows three steps: first some general considerations of innovation and standards, than review of studies of the relation innovation and standardisation and third the development of the five criteria.

3.1 Innovation

The conceptualisation of the role of innovation in development of companies, industries and societies has expanded from mostly focusing on product innovation and secondary process innovation [24] into focusing on a broad range of renewal, newness which is often claimed being far more important to company survival and prosperity [24]. This broader range of innovation encompasses financial innovation, Business model innovation, management and organisational innovation, technological and IT innovation, innovation in networks, alliances and communities [23], [25], service innovations and customer relations innovations, such as channels and brands [24, 26]. Also a lot of interest has been allocated to involving users in innovation [22, 23, 26]. This development complicates defining innovation and understanding how innovation impact on business development. OECD has elaborated the following definition:

"An innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organizational method in business practices, workplace organization or external relations." [27].

This definition is used here in an embedded manner. Innovation processes and innovations and standards as viewed as sociomaterial, with social and material (technical) aspects closely interwoven [18, 28].

3.2 Standardization

A standard can be defined as a specific type of rule meant for common use (by many) [29]. This use would equally often be of voluntary character as sanctioned by some authority. Moreover in a sociomaterial perspective a standard is inseparable from the consensus creating processes needed to create, maintain and develop a standard and the body mandated to facilitate these processes [21]. This is actually also reflected in the International Standardisation Organisation's understanding of standardization (quoted from [29]):

"[a] document, established by consensus, and approved by a recognised body, that provides for common and repeated use, rules, guidelines, or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context"

Moreover in a sociomaterial view processes of arriving at common rules are negotiated and complex and standardization can occur in context where mandated bodies are not involves [30].

3.3 Innovation and Standardization

Early work on innovation and standardization [31, 32], often pointed at the constraints and dysfunctional effects of standardization, where later works overwhelmingly advocate standardization [33-35]. In these works most focus is implicitly on product or process innovation whereas the broader set of possible innovations is rarely treated. Moreover where improved interoperability has been shown to improve company performance, [36] others has measured the losses of poor interoperability in the building sector [9] both of which results add to this indirect effect that improved performance can have on innovation. Importantly the flexibility of standards would also impact on its enabling features vis a vis innovation [34].

Scholars working on mass customization and modularity in product design [37-38], point at first the gains involved in standardizing certain repeated elements of the product structure allowing the design to focus on the specific more customer oriented elements. Such a mass customization strategy fits well with large complex building projects like hospitals that standardized interfaces between subsystems involve. Second modularity of the product design is equally well suited. Again here most focus is on product standardization and secondary process standardization, even if [38] do mention some of the organizational and managerial implications of creating products through mass customization.

It should be noted that building information and its handling in projects tends to be highly volatile and hap hazard because of the large number of players, components and processes [21]. [9] investigates the losses related to poor interoperability amongst the AEC and real estate players.

As a perspective the possibilities of establishing markets beyond single (unique) products and even mass markets would also involve standards as enabler for such market innovation [41]. And such development might lead to the change of entire business models of the involved companies, i.e. to business model innovation [42]. Finally large scope standardizations would encompass sector wide innovation, i.e. systemic innovation [43].

However standardisation continues to involve barriers and pitfalls also in its involvement with innovation. [44] in their empirical study of standards in use, point at the danger of finitism – attempting to create standards covering all aspects of a domain, which risks "locking" the use processes and ultimately leading to non-use of standards because they indeed become perceived as barriers.

Summing up, standards impact on innovation in the following ways

1. Standardization of product and process elements and improved interoperability indirectly provides resources for innovation as the standardized elements require less resources [34, 35]
2. Standardization can enable efficient repetition and the engineering of innovation for single customers, i.e. a mass customization strategy of product development, [38-39]
3. Standardization stabilizes processes in a volatile project based environment [21]

4. Improved interoperability and interfaces between subsystems enable product innovation [40]
5. Standardization enables entry of products into larger markets [41]

Adding to this list but in a more secondary manner innovation in the business model in both the business services organizations [42] and in public organisation, similarly in management, in the financing of the building, such as PPP might also be relevant [1,4, 42,43].

4 The Hospital Project

The present hospital building activity in Denmark is substantial [4]. DNV-Gødstrup is nevertheless one of the largest hospital building projects in Denmark. 130.000 new square meters are designed, constructed and erected over a decade and from 2020 the hospital will be able to make 47.000 operations per year.

Between summer 2012 and autumn 2013, the first major test project of parts of CCS, namely room and building classification, was carried out in this hospital project. The first prototypes and testing activities were developed during the autumn. In this context, the building client became allied with six software suppliers. Together, their six systems cover parts of the information flow from early conceptual design of a building (one system), over detailed design (two CAD-systems and a BIM system), cost and budget calculation (one system), and space management (one system). According to the project manager, the systems are able to identify building components, classify them and sort them. This also involves data flows supported by the chain of the six systems:

"At [the hospital] we are now at classification of rooms and about to classify building components. The six participating IT companies can actually all, almost all, classify. We have made an internal demo of an information flow: [list of the six systems]. The programs are capable of doing that. With CCS we can classify, sort, identify. The programs are further than I thought" (project manager, Nov. 2012).

The classification standard was implemented in six IT systems constituting a common infrastructure and covering important parts of early conceptual design and detailed design.

The central advantage of using CCS is envisioned to be that it integrate the four elements of 1. Classification of building information, 2.property data of building components and rooms in the building, 3. information levels to control the design and further processes, and 4. rules for measuring building components (metrics). It is one common system for handling building information in contrast to a normally completely fragmented building design context. The DNV- project was the first test of these visions.

4.1 Standardisation with Indirect Impact

Implementation and use of standards in the building industry such as CCS implies that a more common terminology and structures are implemented, concerning products, processes and data exchange. This also enabling commencing bridging between the many different IT-solutions that are in use across the companies participating in the project teams and enables smooth communication between the IT systems. A common structuring of buildings by designers is to perceive it as composed in rooms for various purposes, which therefore is one area of classification in CCS. Moreover it is also a demand that the standards should allow new innovative solutions, giving users a tool rather than a strict and detailed coding, that might create barriers for "thinking outside" the given frames, a recurrent initiator of innovation.

To do so CCS is in principle shaped as a collection of terms and concepts that can be brought together in different ways within the code structure. Users have a large room for manoeuvre to specify products and processes at the lower levels in the classification, while CCS also maintains a precise coding structure for specification at the first three to four levels, enabling IT-interpretation of the specification and exchance of data between it-systems.

A classification system is traditionally defined as a hierarchy of classes, and thus the number of classes of components or processes is fixed in the system. E.g. the classification system OmniClass has 211 different classes for doors, and the distinction of the classes primary are based of the properties of doors [45]. By it OmniClass actually determines the number of classes of components, which may exist for the user using this classification system. It can be right a challenge find the right class among the numerous types of classes or even to find a class which is adequate for a new innovative designed door component.

In contrast to traditional hierarchical classification systems like OmniClass, CCS has only one class for all types of door components. But CCS also allows you to add CCS-properties to the class (as many as you wish or need) thus you actually defined your own class-specification for your doors. CCS classification combined with CCS properties gives the users nearly an infinite numbers of combination possibilities to specify exactly the classes of building components needed. A well-defined syntax for code structure for specifying the created classes, act as a digital syntax for digital communication and exchange of the specification of the class.

The flexibility of CCS supports the users to be able to classify new solutions and at the same time be sure of, that the CCS code structure ensures that IT systems are able to interpret the specification og the class. Thus the implementation and the use of CCS indirect release resource from digital implementation to innovate.

The consultant unit of the hospital used CCS first to program rooms and later to organise building components preparing tendering documents. The room programming became more structured because all standard and special rooms entered the same structure enabling a move towards more standard rooms. Thereby the room programming became more efficient than usual for such large and complex building, here handling about 4000 rooms, whereof 80% became classified as standard rooms. This indirectly created resources for the design of the remaining unique rooms. Here the envisaged flexibility of CCS was used and evaluated instrumental. The combination of the classification tables and codes worked. Property data however were not readily available and the design team therefore developed their own properties to enter in the CCS structure.

The client evaluates that better structured building component data has given considerable positive effects in the tendering of contractor contracts, where the design team of engineering and architects are enabled in several ways.

4.2 Standardization Supporting a Mass Customization Platform

CCS is a structure for a digital platform for the building product and supports the notion of a product master [38]. The database-embedded product master, as the CCS server, supports the generic product properties and structures as well as the specific.

The hospital project involves a large amount of repeated components, building elements up to entire blocks of beds. The engineering and architectural design group work with room programming using CCS reduced the number of special rooms and increased the standard room to 80% of the 4000 rooms. The design of the last 20% that could not be standardized as they were unique special rooms. This reveals using mass customisation strategies with CCS for example design, function and equipment such as doors, windows, and HVAC equipment (and for example oxygen) has become more efficient than usual for such large and complex building.

The document analysis of the developed standards supported that the CCS standard might be offering too little depth in the fixed structure and too much flexibility for local appropriation. Such a partition runs counter to obtaining efficiency through the repetition of the core/product master [38, 39]. Some of the building consultants involved in the hospital project expressed similar concerns.

4.3 Standardisation Stabilises Building Processes

It is considerably easier to develop new innovative digital solutions in the building industry, i.e. the building processes from conceptual design, detailed design, production and operation, when your work platform is based on a well-known and stable production environment. A production environment which does not set any technical limitation, but supports the creation of new innovative solutions, is also more or less a necessity for a creative result.

To create a stable production environment, you need to stand on a standardised digital platform, where well-defined and structured building terms are implemented and where simple data operations, like data creation, search, sorting, exchange, can be executed seamless without any IT-specified knowledge for the users.

The purpose of CCS is to create a standard digital platform for the building industry, on which the different parties are able to create a stable digital productions environments. An environment from which they can create new innovative solutions for their building projects, whiteout worrying about whether the solutions can enter into a digital structure or can be communicated and be interpreted digital by other parties' it-systems.

To ensure a stable digital production environment you need to implement the CCS standard platform in your working methods and IT-systems, and surely also upgrade your employees competences in how the this new platform are able to support them in producing new innovative solution based on the ever ongoing demands for new solutions and progress.

At the hospital project CCS has first supported a systematic detailed planning of the design process, digital architecture and work method. This planning created stability as it afterwards had to be followed strictly. Moreover and second CCS supported ”data discipline” in all the sub activities. Also the enabled reuse of the CCS elements supported stabilisation. An important prerequisite for this was a systematic training effort of the members of the design team, especially those involved with Building Information Modelling (BIM).

4.4 Product Development Enabled by Interoperability and Interfaces

Interoperability and interfaces within the product structure is an organised way to enable a number for players to contribute to product innovation through digital design collaboration and communication. In a somewhat similar vein as the well known apple i-phone platform, the CCS platform provides extensive space for supplementary, enlarging innovations “as long” as they comply with the platform they can be taken onboard.

At the hospital project, CCS, enables handling of many types of digital objects, such as documents, BIMs, spreadsheets, data sets (in databases) and drawings. The classification codes enable automatic identification by distinguishing between the items. It is common in large complex buildings project to use considerable resources for coordination of the design activities, which is located in the many participating companies at numerous places and usually involving many different IT –systems and data structuring approaches. Here CCS supports interoperability also by standardising the interfaces between these systems.

4.5 Standardisation Opens Larger Markets for Products

Standardisation of design, production and operation processes is enabled by use of BIM. Process innovation through reuse and iterations are important new opportunities: Reuse of parts of the sustainable design is enabled by well-structured data ordered as objects and more feasible if one encounter a (larger) Nordic market and markets beyond that. Several building consultants involved in the test project have won new project in the health care sector.

IT suppliers participating in the hospital project have used their experiences with CCS to incorporate the classification in their IT-systems. They are currently marketing that in the Nordic and Baltic region. More in general there is a large global market for hospital design providers, where experiences of CCS can be transformed into design service offerings globally. Here it is likely however that competing standards will be part and parcel of the future market conditions [10, 11, 42].

5 Discussion

In the following the five innovation –standard relation types is first discussed, also looking at enablers and barriers within each dimension. Then an overall discussion is summing up the results.

The *first* dimension is the indirect enabling. The standardisation is expected to indirectly provide resources for innovation. The case shows how standardisation in the tendering process prepares for cost reductions that indirectly can provide more space for innovative solutions developed in the products. The client thus claims to have saved 2, 68 mill. Euro in this way [14]. The communicated tendering material is easier to access, better structured, and more homogenous, which in turn generate more comparable and cost efficient tenders. However the implementation and co-testing of the standards with the development organisation also required substantial investment in terms of hours and human resources. One issue being that the testing of the standard commenced before the standard was fully developed, i.e. a beta version was provided in the beginning and later improved.

Second standardisation can promote efficient repetition and the engineering of innovation for single customers (i.e. a mass customisation). The room programming and the standardisation of it gave considerable result enabling innovation both on the standard side and the customer specific side. Also here it required considerable implementation work to reach this result in the test project.

Third standardisation is expected to stabilize processes. In the hospital project this occurred as improved planning and also a relatively strict practice of following these plans providing stabilisation. It was carried out in a manner that demanded systematic change management, provided through training.

Fourth an improved interoperability and interfaces between subsystems enable product innovation. The classification enabled handling of many types of digital objects, and supported interoperability also by standardising the interfaces between these involved systems.

Fifth standardisation is expected to create larger markets for products, which in the hospital project both occurred as marketing of new solution and a still unexploited potential for offering new services. The IT-suppliers were provided with competences after having developed relevant IT-tools, which enable them diffusing the standard. More specifically the public building client evaluates that better structured data on building component give considerable positive effect in the tendering of contractor contracts, where the design team of engineering and architects are enabled i several ways by the computerized information standard.

As the standard CCS at present has only been tested in Denmark, its diffusion will face technical and social barriers. Technically CCS would have to coexist with a range of other building information standards [11], and the interface to these other standards in Scandinavia and beyond are important rather than imagining that CCS would creatively destruct the other standards by technical superiousity. As noted by [12, 13] multiple standards in domains are a likely future scenario. Importantly the standards in use in a Scandinavian setting are overlapping and partially mutually compliant [47]. CCS builds on the international ISO standard 13006-2, which broadly specifies a systems model of building processes organising them as resources as input to a process (transformation) that lead to a result and that are feeded with properties of the involved components and entities. This and other commonalities do not overcome the technical differences which are also coupled to social aspects such as the organisation of influence on the further development of standard. It has previously been seen and it is likely that the trajectory of CCS will involve a hybridization of the standard, i. e. a mixing of content, with other Scandinavian standards [48].

The range of possible future innovation involves financial, organisational and managerial innovations as well as community innovation once the hospital is built [4, 23]. Importantly the IT supported operation of the built facilities would be enabled by a comprehensive “As-built” building information model. It would be a possibility to do space management, technical maintenance, reconfiguring of the use and more in the continual further development of the hospital facilities.

The analysis shows that effects of standards impacts on a series of players in the industry and the effects for each party are often counter to each other. The effects are intraorganisatoric with for example classification of rooms that enable designers overview of the complex building. But also the interorganisational effects are very import as when room programming is carried out in interaction with client representatives as well as other designers. Evaluating the impact of standards in test projects can be questionized. Especially during test projects there is a risk that some parties experience they carry out unpaid work for other parties. The test project experiences could be seen as “childrens diseases”, but it should be noted that even if most would think of standards as stable and even rigid, thereby viewing an implementation as a short barrier to overcome before a long stability, this is hardly the contemporary case. A series of dynamics would imply that the standard would be under continual development, related to adoption in and between organisations, related to technological development of building materials and methods, related to the promotion, revision and distribution and related to the rapport with or integration into public regulation [29]. Therefore swift implementation would be a demand for a contemporary standard as would the flexibility to adapt to context and change.

6 Conclusion

The impact of building information standards on innovation is both inter- and intraorganisational, contradictory and dynamic. In our case there are two main indirect enablers of innovation; first the building client claims to have saved money, and second the IT-suppliers and design companies appear to have been provided with competences. But a range of other effects occur as well: The public building client evaluates that better structured data on building components give considerable positive effect in the tendering of contractor contracts, where the design team of engineering and architects are enabled in several ways by the computerized information standard. The standardized building information appear to provide cost reduction and stabilised processes that indirectly opens for innovation, but also the standardisation of rooms led to more direct innovation in the design of rooms both in the standardised and in the unique rooms, in line with a mass customisation strategy. Moreover the use of one common standard enables a far better coordination than previously inside a hospital project like this one, even if the broader and longer term perspective probably involves handling more standards over time and place and/or an increasing hybridisation of standards, i.e. they would tend to overlap and with some common structures and components. Finally we currently witness IT-suppliers using their experience with classification in providing new solutions and marketing these on a Nordic market, whereas we still have to see the event of the design companies following the same path of providing new services related to improved classification.

References

1. WHO. Investing in hospitals of the future. World Health Organization, Copenhagen (2009)
2. OECD: Sweden 2013 Raising Standards. OECD Reviews of Health Care Quality. OECD (2013)
3. Hågøy, T.: Nye helsebygg for 73 milliarder på 10 år. Dagensmedisin (2013), `http://www.dagensmedisin.no`
4. Danske Regioner. Nye sygehusbyggerier: Det erhvervsmæssige potentiale. Danske Regioner, København (2011)
5. STD: Sector Review The Consulting Engineering and Architectural Groups. A Swedish and International survey. The Swedish Federation of Consulting Engineers and Architects (STD), Stockholm (2012)
6. Bernstein, H.M.: The Business Value of BIM in Europe. McGraw-Hill, Bedford (2010)
7. Bernstein, H.M.: The Business Value of BIM in North America. McGraw-Hill, Bedford (2012)
8. Eastman, C., Teicholz, P., Sacks, R., Liston, K.: BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors, 2nd edn. Wiley & Sons, Hoboken (2011)
9. Gallaher, M.P., O'Connor, A.C., Dettbarn Jr., J.L., Gilday, L.T.: Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry. National Institute of Standards and Technology, Gaithersburg (2004)
10. Laakso, M., Kiviniemi, A.: The IFC Standard - A Review of History, Development, and Standardization. Journal of Information Technology in Construction 17, 134–161 (2012)
11. Cuneco: Website in English of cuneco, the Danish center for classification and productivity in construction (accessed January 2014)
12. Berman, D.: Make It What Way? The Impact of Multiple Standards Regimes. Int. Jrnl. of Soc. of Agr. & Food 20(1), 69–89 (2013)
13. Reinecke, J., Manning, S., von Hagen, O.: The Emergence of a Standards Market: Multiplicity of Sustainability Standards in the Global Coffee Industry. Organization Studies 33(5-6), 791–814 (2012)
14. Andersen, U.: Digitale tilbudslister sparer Gødstrup Sygehus for mindst 20 mio. kroner. Ingeniøren (September 6, 2013)
15. EU: Directive 2012/27/EU on energyefficiency. The European Parliament and of the Council. Bruxelles (2012)
16. Walsham, G.: Interpreting Information Systems in Organizations. Wiley, Chichester (1993)

17. Bryman, A., Bell, E.: Business Research Methods, 3rd edn. Oxford University Press, Oxford (2011)
18. Orlikowski, W.J., Scott, S.V.: Sociomateriality: Challenging the separation of technology, work and organization. Annals of the Academy of Management 2, 433–474 (2008)
19. Webster, J., Watson, R.T.: Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly 26(2), 13–23 (2002)
20. Greenhalgh, T., Peacock, R.: Effectiveness and Efficiency of Search Methods in Systematic Reviews of Complex Evidence: Audit of Primary Sources. British Journal of Medicine 331(7524), 1064–1065 (2005)
21. Koch, C., Jacobsen, K.: Computerized Information Standards Enabling Innovation in Public Procurement of buildings. In: Proceedings Human Computer Interaction, Creta Maris, Heraklion, Crete, Greece, June 22-27 (2014)
22. Koch, C., Chan, P.: Projecting an Infrastructure - Shaping a community. In: Carrillo, P., Chinowsky, P. (eds.) Proceedings EPOC 2013 Conference. Engineering Project Organization Society, Colorado (2013)
23. Botero, A., Vihavainen, S., Karhu, K.: From closed to open to what?: an exploration on community innovation principles. In: Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era (2009)
24. Fagerberg, J., Mowery, D.C., Nelson, R.R.: The Oxford Handbook of Innovations. Oxford University Press, Oxford (2004)
25. Koch, C.: Innovation networking between stability and political dynamics. Technovation 24(9), 729–739 (2004)
26. Chang, T.-R., Kaasinen, E.: Three user-driven innovation methods for co-creating cloud services. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011, Part IV. LNCS, vol. 6949, pp. 66–83. Springer, Heidelberg (2011)
27. OECD: Oslo Manual, Guidelines for collecting and interpreting innovation data, 3rd edn. OECD Publishing, Paris (2005)
28. Timmermans, S., Epstein, S.: A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. Annual Review of Sociology 36, 69–89 (2010)
29. Brunsson, N., Rasche, A., Seidl, D.: The Dynamics of Standardization: Three Perspectives on Standards in Organization. Organization Studies 33(5-6), 613–632 (2012)
30. Hanseth, O., Jacucci, E., Grisot, M., Aanestad, M.: Reflexive Standardization: Side Effects and Complexity in Standard Making. MIS Quarterly 30(2), 563–581 (2006)
31. Farrell, J., Saloner, G.: Standardization, Compatibility, and Innovation. The RAND Journal of Economics 16(1), 70–83 (1985)
32. Kahan, M., Klaussner, M.: Standardization and Innovation in Corporate Contracting (Or "The Economics of Boilerplate"). Virginia Law Review 83(4), 713–770 (1997)
33. Ekholm, A., Blom, H., Eckerberg, K., Löwnertz, K., Tarandi, V.: BIM – Standardiseringsbehov. Svenska Byggbranschens Utvecklingsfond, Stockholm (2013)
34. Lillrank, P.: The Quality of Standard, Routine and Nonroutine Processes. Organization Studies 24(2), 215–233 (2003)
35. Sandholtz, K.: Making Standards Stick: A Theory of Coupled vs. Decoupled Compliance. Organization Studies 33(5-6), 655–679 (2012)
36. Zhao, K., Xia, M.: Forming Interoperability Through Interorganizational Systems Standards. Journal of Management Information Systems 30(4), 269–298 (2014)

37. Hanseth, O., Nielsen, P.: Infrastructural Innovation: Flexibility, Generativity and the Mobile Internet. International Journal of IT Standards and Standardization Research (IJITSR) 11(1), 27–45 (2013)
38. Piller, F.T., Tseng, M.M. (eds.): Handbook of Research in Mass Customisation and Personalization. World Scientific, New Jersey (2010)
39. Poorang, A., Piroozfar, E., Piller, F.T.: Mass Customisation and Personalisation in Architecture and Construction. Routledge, London (2013)
40. Clark, K.B., Baldwin, C.Y.: Design Rules. The Power of Modularity, vol. 1. MIT Press, Cambridge (2000)
41. Schilling, M.: Strategic Management of Technological Innovation. McGraw-Hill, New York (2008)
42. Teece, D.J.: Business Models, Business Strategy and Innovation. Long Range Planning 43(2-3), 172–194 (2010)
43. Malerba, F.: Sectoral systems of innovation: concepts, issues and analyses of six major sectors in Europe. Cambridge University Press (2004)
44. Bowker, G.C., Star, S.L.: Sorting Things Out: Classification and Its Consequences. MIT Press, Cambridge (1999)
45. Omniclass: Omniclass website, `http://www.omniclass.org/` (accessed at January 2014)
46. Fomin, V.V.: Standards as Hybrids: An Essay on Tensions and Juxtapositions in Contemporary Standardization. International Journal of IT Standards and Standardization Research 10(2), 59–68 (2012)
47. Ekholm, L.: Häggström Mappning mellan systemen för byggklassifikation BSAB 96 och DBK Lund University. In: Lund 2011 (2011)
48. Shen, X., Graham, I., Stewart, J., Williams, R.: Standards Development as Hybridization. International Journal of IT Standards and Standardization Research (IJITSR) 11(2), 34–45 (2013)

Extending a Metamodel for Formalization of Data Warehouse Requirements

Natalija Kozmina and Laila Niedrite

Faculty of Computing, University of Latvia, Raina blvd. 19, Riga
{natalija.kozmina,laila.niedrite}@lu.lv

Abstract. In performance measurement systems that are built on top of a data warehouse, the information requirements in natural language are different performance indicators that should be stored and analyzed. We use the requirement formalization metamodel to create a formal requirement repository out of information requirements in natural language. In the course of this research we tested the compatibility of the existing requirement formalization metamodel applying it to a set of over 150 requirements for the currently operating data warehouse project. As a result, we extended the formal specification of information requirements with some additional classes like themes, grouping, and requirement priorities, and relationships discovered in this case study. We discussed benefits of requirement prioritization and advantages of transferring requirement priority values to schema elements with and aim of detecting schema elements to be incorporated into dashboards.

Keywords: data warehouse, performance indicators, requirement formalization, prioritization.

1 Introduction

Companies use performance measurement systems to control their progress. Measurement results are compared to target values to evaluate whether companies' goals are achieved. If companies already use a data warehouse for analytical purposes, they can also be used as a foundation for a Performance Measurement System.

"A data warehouse is a subject-oriented, integrated, non-volatile, and time-variant collection of data in support of management decisions" [1]. A data warehouse stores data according to a multidimensional data model, which should be built in compliance with the analysis requirements of a company. Therefore, we can speak about the information requirements [2] for data warehouses that determine two types of data to be modeled – quantifying and qualifying data, and elements of multidimensional paradigm, e.g. dimensions, hierarchies, and cubes.

There exist different methods to construct conceptual models for data warehouses, which can be classified as supply-driven or demand-driven. Supply-driven methods determine the existing information requirements during the analysis of data models of

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 362–374, 2014.

data sources and more or less automated transform into the data warehouse model. The limitation of supply-driven approach is that the real analysis needs may not be found out due to the operational nature of data warehouse's data sources and their data.

In demand-driven approaches the information needs are gained by interviewing users, therefore, the conceptual model of data warehouse depends on how precise the users formulate and data warehouse developers formalize the analysis needs. Precisely documented information requirements may serve further as a basis for semi-automated methods for development of a conceptual model of a data warehouse that afterwards can be checked for existence of source data.

In case of performance measurement systems that are built on top of a data warehouse we can consider the information requirements as different performance indicators that should be stored and analyzed. In our previous works we have proposed a formal specification of indicators [3] and a method for transforming formally expressed information requirements or indicators into a conceptual model of a data warehouse [4]. The formal specification of indicators was built after analyzing a set of indicators from indicator database [5].

The goal of our research presented in this paper is to use the findings in previous research [3] in real data warehouse project to extend the formal specification of indicators with elements discovered in this case study. Afterwards, these changes should be reflected also in the method for semi-automated data warehouse schema construction [4].

The rest of the paper is organized as follows. Related work and background is presented in Section 2. Section 3 describes the improvements of the requirement formalization metamodel followed by an example. A short overview of requirement prioritization techniques and benefits of setting priorities is given in Section 4. Section 5 finalizes the paper with conclusions and future work.

2 Related Work and Background

Performance indicators should be aligned with the strategy of the company. During the elicitation of requirements for a performance measurement data warehouse the information needs are expressed as more or less complex sentences that describe how the performance can be measured. In this case, these sentences or requirements represent performance indicators. The indicators can be defined on various levels of formality.

2.1 Related Work

Some research on how to specify performance indicators has been done and is described in [6, 7, 8].

The authors of [6] propose a formal language for modeling goals based on performance indicators. Goal satisfaction could be controlled and evaluation of organizational performance could be performed.

In [7] the authors propose a formal language for indicator definition by introducing the sorts of indicators, predicates and functions included in it. Relationships between indicators are defined. The authors argue that the usage of the considered specification language can be informal, semi-formal, graphical or formal. They argue that the requirements can be reformulated from natural language expressions to more formal. However, they do not use the formal representation of the indicator as an essential part of their specification language.

In [8] the ideas of [7] are extended and the formalized indicators are integrated with other performance-oriented concepts, which describe actual performance measurement and evaluation, e.g. processes, goals, agents, etc. Formalizing performance indicators [7] includes the definition of all relevant characteristics of indicators, e.g. name, definition, type, timeframe, etc.

The authors of the research presented in [9] have based their proposal on User requirement notation [10], which is a standard used mostly in telecommunication field. It describes concepts that can be also applied in business process modeling, including goal and requirement modeling. In [9] they integrate the KPI concepts with URN metamodel. Such indicator features as target value, threshold value, worst value, and others are defined.

Our approach formalizes not only different features of indicators like name, type, etc., but also tries to decompose the indicator definition in more detailed parts according to proposed indicator definition metamodel. Also, the goal for such formalization is not only to describe the usage aspects of indicators in performance measurement and their dependencies. We will use the formalized indicator definition to semi-automatically generate a data warehouse model, where appropriate data for performance measurement will be stored. We add also some data warehouse development specific features to the model, for instance, requirement priorities.

2.2 Background

In [4] we propose a method for transforming information requirements to the conceptual model of a data warehouse. The main components of this method are depicted in Figure 1.

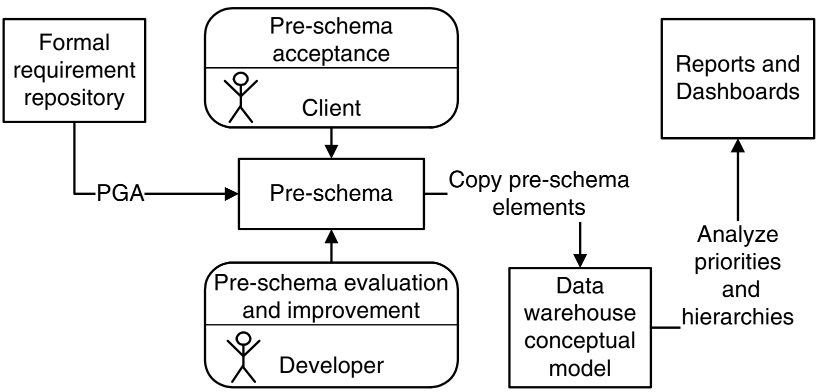


Fig. 1. Pre-schema generation and restructuring (according to [4])

The method uses a set of requirements, which are formalized according to requirement formalization metamodel and stored in the formal requirements repository, and generates a simplified data warehouse schema – a pre-schema or a candidate schema – by the Pre-schema Generation Algorithm (PGA) that analyses the structure of requirements.

On the next stage of the method semi-automated candidate schemas are processed and restructured by developer to remove duplicates and build dimension hierarchies.

The improved schemas can be used as data warehouse schema metadata. All generated pre-schemas are being shown to the client during an interview, where the client should make a decision and choose one pre-schema that meets the requirements for a new schema best of all. The elements of the chosen pre-schema are being copied to the conceptual model of the data warehouse.

In this paper we add requirement priorities to the requirement formalization metamodel. We take advantage of priorities and build requirement hierarchies with a purpose to distinguish schema elements that should be incorporated into dashboards (see more details in Section 4.2).

3 Requirement Formalization Metamodel and Its Application

The previous version of the requirement formalization metamodel was published in [3], and was tested (i.e. used to create formalized requirements) on approximately 330 different indicators (listed in [5]) from such measurement perspectives as customer focus, environment & community, employee satisfaction, finance, internal process, and learning & growth. The most complicated example of an indicator like this would contain a ratio, for instance, a summary information on the percentage of IT expense of total administrative expense in a year would be formally written as "(sum (expense) where expense type = 'IT') / (sum (expense))".

The previous version of the requirement formalization model was tested on a large set of indicators from the business sphere listed in [5]. The goal of our current research is to check the compatibility of the requirement formalization metamodel when applying it to a set of requirements for a real data warehouse project, and extend the formal specification of indicators with elements discovered in this case study.

Then, the same requirement formalization metamodel has been tested on a set of requirements for the currently operating data warehouse of the University of Latvia. This data warehouse accumulates data on student enrolment statistics, student and academic staff activity in e-learning system, strategic indicators, staff workload statistics, etc. The overall number of requirements is over 150. While testing, it was stated that the metamodel should be extended with some additional classes like themes, grouping, and priorities, as well as relationships between classes should be reviewed.

Also, a small part of these requirements were more complex and consisted not only of ratios, but also an evaluation of these ratios (such as "the number of post-docs in elected positions should increase by 10% by next year"), which made us extend and restructure the requirement formalization metamodel (a detailed explanation will be given in Section 3.1).

3.1 Description of the Extended Requirement Formalization Metamodel

The information needs expressed as more or less complex sentences or requirements that describe how the performance can be measured represent performance indicators. Since most of the requirements have a common pattern, we apply a metamodel to reformulate these requirements in a formal way. The metamodel is designed using UML class diagram notation (Figure 2).

Previous Version. In the previous version of the requirement formalization metamodel [3] a Requirement is classified either as Simple or Complex. A complex requirement is composed of two or more requirements joined with an Arithmetical Operator. A simple requirement may consist of an Operation that denotes a command applied to an Object, and an optional Typified Condition. In its turn, an object is either an instance of Quantifying data (measurements) or Qualifying data (properties of measurements) depending on the requirement. A Complex Operation consists of two or more Actions, which are of two possible kinds: Aggregation ("roll-up"; for calculation and grouping) and Refinement ("drill-down"; for information selection). Information refinement is divided into showing details (selecting information about one or more objects), or showing details restricted with a constraint defined by Typified Condition (slicing). Just like requirements, Conditions and Expressions are either Simple or Complex. Complex condition joins two or more conditions with a Logical Operator. A simple condition, for instance, "year > 2013", consists of a Comparison of two Expressions. A complex expression contains two or more expressions with an arithmetical operator in between, whereas a simple expression belongs either to qualifying data (e.g. "year") or to Constants (e.g. "2013").

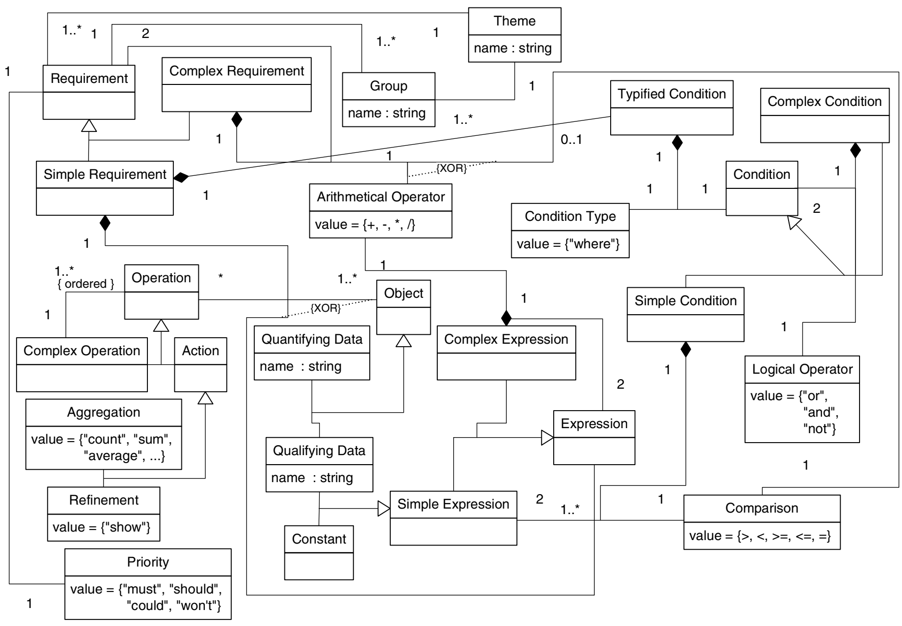


Fig. 2. Extended version of the requirement formalization metamodel

Extended Version. The extended metamodel maintains all existing classes from the previous version. We added the following improvements:

- Each requirement has to have its Priority value (must, should, could or won't);
- Two or more requirements, which make up a complex requirement, may be joined with either an Arithmetical Operator or a Comparison (see a comparison between two complex requirements in Figure 3);
- A simple requirement now may consist of an Operation that denotes a command applied to an Object, or of an Expression (see a constant = "10%" in Figure 3), and an optional Typified Condition;
- Each requirement may refer to one or multiple Groups (e.g. Dynamics, Master studies, Doctoral studies, etc.), whereas a Theme as a coarser level of grouping (e.g. Finance, Education, Customer Focus, etc.) may unite one or more groups. Grouping requirements is needed for several reasons:
  a. To reduce the number of repeating elements in requirements, thus, making them more compact – for instance, if a number of requirements contains one and the same time frame (e.g. year), it can be written just once as a simple requirements "show year";
  b. To unite multiple requirements logically, which would be the natural grouping of reports to be developed later on (e.g. Dynamics, E-learning, Staff statistics, Student statistics, etc.).

3.2 An Example of a Formalized Requirement

Let's consider an example of application of the requirement formalization metamodel (Figure 3). Priority of the following requirement is "could", Theme is "Education", and Group is "Master studies". As these 3 classes solely characterize the requirement, but are not connected to other classes that help to form the requirement, they are excluded from the example in Figure 3.

Informally, the requirement goes as follows: "The ratio of master level graduates in the University of Latvia in 2013, who are employers, has to be 10% of master level graduates in the University of Latvia in 2012". In its turn, it is reformulated this way: "((count (graduate occurrence) where level = 'master' and year = '2013' and status = 'employer') / (count (graduate occurrence) where level = 'master' and year = '2013')) = (10% * (count (graduate occurrence) where level = 'master' and year = '2012'))".

The left column is filled with parts of the requirement statement and all the rest columns (left to right) contain class names of the requirement formalization metamodel.

Here 2 principles of requirement (indicator) reformulation are applied:

- A component to be measured is treated as an aggregated number of all occurrences of this component. For example, "sessions" is reformulated to "count (session occurrence)", where count is the most suitable aggregate function;
- If there are such components as "%", "percent", "percentage", or "ratio", then % is substituted by division of partial quantity by total quantity. For example, "IT expense as a % of total expense" is reformulated to "sum (IT expense) / sum (expense)".

See the full list of principles of requirement (indicator) reformulation in [3]. These principles appeared from our practical experience and serve to translate the requirements from natural language to a state that is compatible with the requirement formalization model.

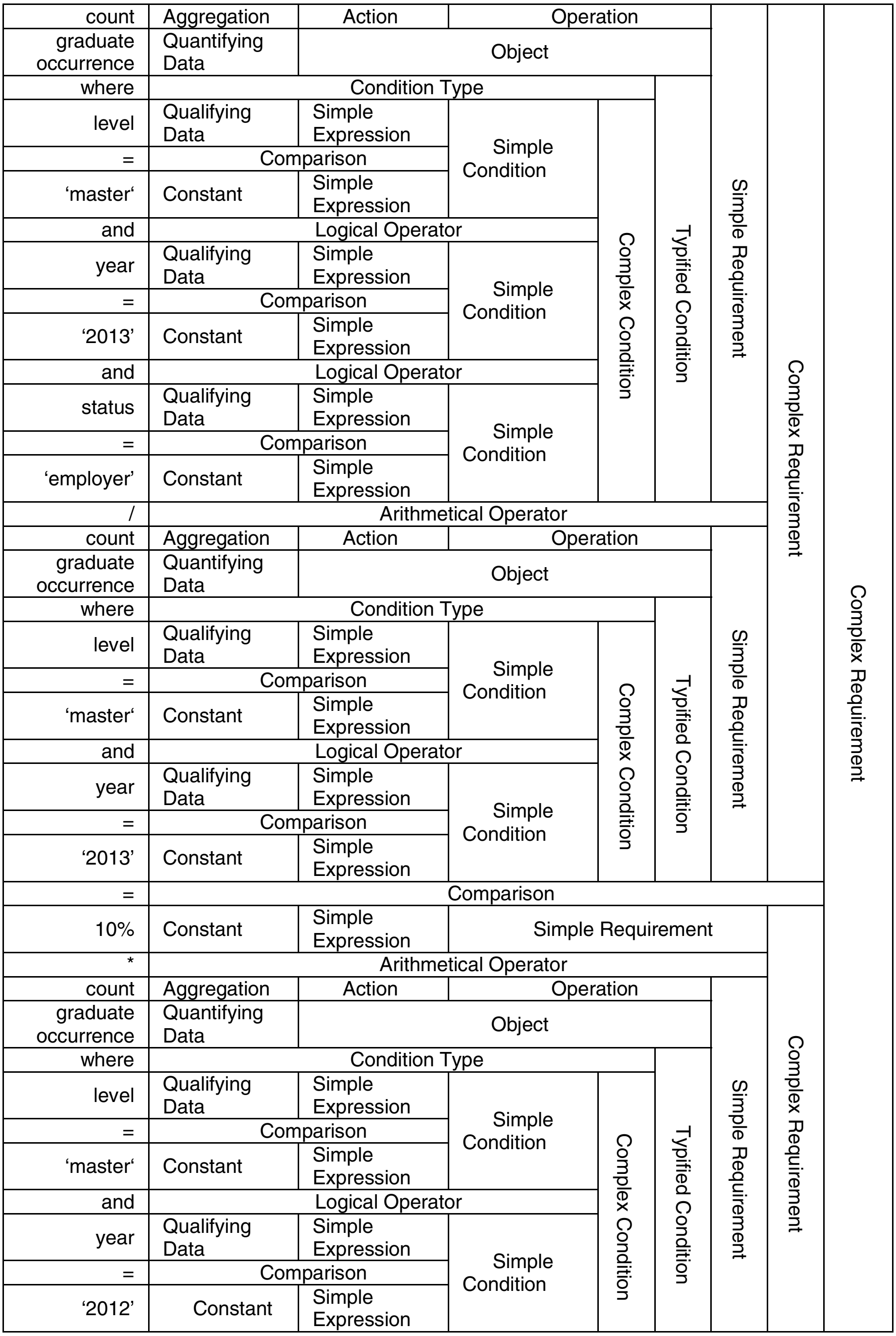


Fig. 3. An example of a formalized requirement

This requirement has a sophisticated structure and it is a complex requirement that consists of two others. We interpreted "has to" as a request for equality, thus, these two complex requirements are linked with "=" sign.

4 Requirement Prioritization

In this Section we give a succinct overview of existing methods for setting priorities to requirements, as well as share our viewpoint on benefits of using priorities in requirement formalization process.

4.1 Techniques for Setting Requirement Priorities

Here a short summary of such techniques as Top-10 Requirements, Ranking, Numerical Assignment (Grouping), 100-Dollar Test, Analytical Hierarchical Process (AHP), and MoSCoW is presented in Table 1.

Table 1. Short descriptions of techniques for setting requirement priorities

Approach	Short Description	Difficulty level
Top-10 Requirements	each decision-maker selects his/her top-10 requirements that are considered more important from a set of requirements [11, 12]	extremely easy [11]
Ranking	each of requirements is assigned a number from 1 to n, where n is the total number of requirements, and all ranks are unique [11, 13]; ranks can be obtained by some algorithm such as bubble sort, etc.	easy [11]
Numerical Assignment (Grouping)	developed by J. W. Brackett in 1990 [14]; the requirements are split into different priority groups depending on its importance (e.g. "mandatory", "very important", "rather important", "not important", and "doesn't matter" [15], although, other meaningful names of the priority groups are acceptable too, e.g. essential, conditional, and optional)	very easy [11]
100-Dollar Test (Cumulative Voting)	it is a prioritization technique where decision-makers are given 100 arbitrary units (e.g. money, working hours, etc.) to share between the requirements [11, 17]; the higher is the amount, the higher is the priority	complex [11]
Analytical Hierarchical Process (AHP)	developed by T. L. Saaty in 1980 [18]; AHP [11, 18, 19] is "a systematic approach for decision-making that involves the consideration of multiple criteria by structuring them in a hierarchical model" [20]	very complex [11]
MoSCoW	it is a requirement prioritization technique to be easily used during requirements analysis and specification; requirements are divided into four groups: must, should, could, and won't according to Business Analysis Body Of Knowledge (BABOK) Guide [21]	very easy

The drawback of **Top-10 Requirements** method might be the difference in requirements that compose a top-10 requirements set of each decision-maker. Whose top-10 requirement set should be considered of a higher priority? Is it possible to satisfy needs of all decision-makers? As a solution one may consider satisfying a subset of the most significant requirements of each decision-maker, thus, mitigating a potential conflict.

Ranking technique is considered to be more appropriate for a single decision-maker, however, in case of multiple participants a mean priority value for each requirement should be calculated. This may produce similar values for a number of requirements, thus, making pointless the prioritization process itself, and may be treated as a drawback of this method.

In **Numerical Assignment (Grouping)** it is crucial to clarify the name of each group, so that each of the decision-makers would have the same understanding of its definition. To keep the groups balanced, it is recommended to put restrictions on the total number of requirements in each group (e.g. more or equal to 25% in each group) [16].

100-Dollar Test (Cumulative Voting) method is not suitable for a large number of requirements (e.g. if there are 50 requirements, then on average there are 2 points to assign to each requirement); however, the number of units may be extended to 1000 or more to deal with this issue. Another potential drawback of this method is that it requires attention when distributing or re-assigning priority points, so that the total sum of points would be the same.

Analytical Hierarchical Process (AHP) is conducted by comparing all possible pairs of hierarchically classified requirements to define which of these requirements has a higher priority. Since the comparisons are performed pairwise, there are n*(n-1)/2 of them at each level (n is the number of requirements). The drawback of this method is that the larger is the number of requirements, the more significantly increases the number of comparisons.

In **MoSCoW** analysis requirements are divided into four groups: must, should, could, and won't, which are defined in Business Analysis Body Of Knowledge (BABOK) Guide [21] as follows:

- "Must" describes a requirement that must be satisfied in the final solution for the solution to be considered a success;
- "Should" represents a high-priority item that should be included in the solution if it is possible; this is often a critical requirement but one that can be satisfied in other ways if strictly necessary;
- "Could" describes a requirement, which is considered desirable but not necessary, and will be included if time and resources permit;
- "Won't" represents a requirement that stakeholders have agreed will not be implemented in a given release, but may be considered for the future.

Requirement priorities may be redefined when needed, however, it is advised to consider the proportion of maximum total effort: must – 60%, should – 20%, could – 20% (won't requirements are not included into it). MoSCoW analysis works best when priorities are discussed and assigned in groups. MoSCoW is not evaluated in [11], but since it is similar to numerical assignment, we evaluate it as “very easy”.

We have chosen MoSCoW analysis as a requirement prioritization technique, because we were interested in a fast and straightforward approach that doesn’t require complex calculations during re-prioritisation process, and that works best for assigning priorities in small groups of decision-makers (1-5 people). Comparing to Numerical Assignment (Grouping), in MoSCoW analysis the number of priority groups is strictly limited to 4, which cannot be altered, and definitions of each priority group are given in BABOK guide [21].

4.2 Benefits of Requirement Prioritization

In our case, as requirement elements are tightly connected to candidate schema elements, setting priorities to requirements would help to answer the following questions:

- Which of the planned reports should be developed prior to others?
- How requirement priority values are propagated to schema elements?
- Which schema elements to incorporate into dashboards?
- Which candidate schema is most likely to be accepted by client?

Let’s take a look at each of the above-mentioned points in more detail.

Which of the planned reports should be developed prior to others? It is worthy to notice that the structure of a formalized requirement (i.e. the one that includes qualifying and/or quantifying data with or without additional restrictions) is such that it allows us to build a data warehouse report containung schema elements that correspond to qualifying and quantifying data in requirements. In other words, it is quite an easy task to define the potential reports out of initial requirements stated by client.

Thus, having split all the requirements into 4 groups – i.e. must, should, could, won’t – it is possible to create exactly 4 groups of labels for reports respectively in the context of time – namely, most urgent, urgent, less urgent, not urgent. This approach would help to sort the report that should be created prior to others.

How requirement priority values are propagated to schema elements? We described a method of transforming requirements to the conceptual model of a data warehouse (candidate schemas or pre-schemas) in one of our previous papers [4]. A pre-schema generation algorithm (PGA) is employed for distinguishing data warehouse schema elements in formalized requirements, which are stored in formal requirement repository. Thus, if there is some requirement R with a certain priority P, then all shema elements derived from the requirement R (i.e. measures and attributes) have their priority value set to P. Imagine that one and the same schema element (e.g. a Study Program attribute) has more than one priority value (e.g. must, could) gained from a set of requirements with various priorities. If a schema element has multiple priority values, then the one of the higher value is assigned (e.g. a Study Program attribute is assigned a "must" priority value).

Which schema elements to incorporate into dashboards? Dashboards often demanded by decision-makers should not be overwhelmed with data. Only the most essential and summary data is represented in dashboards.

Therefore, our goal is (i) to detect schema elements with highest priorities from the corresponding requirements, and afterwards (ii) to check if any of these elements build up data hierarchies.

The next step would be the creation of requirement hierarchies based on hierarchies in schema elements. The PDA described in [4] may determine attributes and measures from requirement objects, i.e. qualifying and quantifying data respectively. It means that we can analyze requirements that contain the same quantifying data (corresponding to measures) and typified conditions, but different qualifying data (corresponding to attributes).

Suppose, we have a pair of (already formalized) requirements such as:
R1: show course count (user session occurrence) where user role = "student"
R2: show course category count (user session occurrence) where user role = "student"
Consider a Course hierarchy in Course dimension: Course → Course Category. Here "user session occurrence" in e-learning system is related to a measure, whereas "course" and "course category" are related to attributes. In this case, R1 → R2 is a requirement hierarchy example, because corresponding schema elements form a hierarchy. A dashboard report in this case would be the one based on R2 requirement.

Finally, in a given candidate schema we select schema elements that are related to the requirements of the coarser level of granularity (e.g. R2) with highest priority and treat these schema elements as components of a potential report for a dashboard. Dashboard reports may be explored more in-depth sliding down to finer levels of granularity of one or another axis.

Which candidate schema is most likely to be accepted by client? A candidate schema, which includes the largest number of schema elements corresponding to components of requirements with higher priority (i.e. must, should), is the one that is most preferred by the client. There may be more than one way to evaluate each candidate schema; however, the most natural way is to count schema elements of each priority value and sort the acquired 4 values by priorities (must, should, could, and won't) in descending order. Thus, one may obtain a sorted list of candidate schemas based on requirement priorities.

5 Conclusions and Future Work

In this paper we reviewed the requirement formalization metamodel, which is necessary for creating a formal requirement repository out of information requirements in natural language. This metamodel is one of the constituent parts of the methodology for transforming requirements into a conceptual model of the data warehouse [4].

We conducted a case study that consisted of testing the existing requirement formalization metamodel, i.e. the findings in our previous research [3], on a set of requirements for a real currently operating data warehouse project of the University of Latvia. These requirements related to student enrolment statistics, student and academic staff activity in e-learning system, strategic indicators, staff workload statistics, etc. The overall number of requirements was over 150.

Due to a specific structure of requirements that contain an evaluation of ratios (such as “the number of post-docs in elected positions should increase by 10% by next year”), it was stated that the metamodel had to be restructured and extended with some additional classes like themes, grouping, and requirement priorities, as well as relationships between classes had to be reviewed.

Having chosen MoSCoW analysis as the most suitable requirement prioritization technique, we addressed such questions as (i) which of the planned reports should be developed prior to others, (ii) how requirement priority values are propagated to schema elements, (iii) which schema elements to incorporate into dashboards, and (iv) which candidate schema is most likely to be accepted by client.

Our future work would include adapting the Pre-Schema Generation Algorithm (PGA), which is a part of the metodology for semi-automated data warehouse schema construction and requirement implementation model (both presented in [4]), to the changes introduced to the requirement formalization metamodel in terms of this paper.

Acknowledgments. This work has been supported by the European Social Fund within the project “Support for Doctoral Studies at University of Latvia”.

References

1. Inmon, W.H.: Building the Data Warehouse, 3rd edn. Wiley Computer Publishing (2002)
2. Winter, R., Strauch, B.: A Method for Demand-driven Information Requirements Analysis in Data Warehousing Projects. In: Proceedings of 36th Annual Hawaii International Conference on System Sciences (HICSS 2003), pp. 1359–1365. IEEE, Waikoloa (2003)
3. Niedritis, A., Niedrite, L., Kozmina, N.: Performance Measurement Framework with Formal Indicator Definitions. In: Grabis, J., Kirikova, M. (eds.) BIR 2011. LNBIP, vol. 90, pp. 44–58. Springer, Heidelberg (2011)
4. Kozmina, N., Niedrite, L., Golubs, M.: Deriving the Conceptual Model of a Data Warehouse from Information Requirements. In: Proceedings of the 15th International Conference on Enterprise Information Systems (ICEIS 2013), Angers, France, vol. 1, pp. 136–144 (2013)
5. Parmenter, D.: Key Performance Indicators: Developing, Implementing, and Using Winning KPIs, 2nd edn. Jon Wiley & Sons, Inc. (2010)
6. Frank, U., Heise, D., Kattenstroth, H., Schauer, H.: Designing and Utilizing Business Indicator Systems within Enterprise Models - Outline of a Method. In: Loos, P., Nüttgens, M., Turowski, K., Werth, D. (eds.) Modellierung betrieblicher Informationssysteme (MobIS 2008) - Modellierung zwischen SOA und Compliance Management, vol. 141, pp. 89–105. LNI, Saarbrücken (2008)

7. Popova, V., Treur, J.: A Specification Language for Organizational Performance Indicators. Applied Intelligence Journal 27(3), 291–301 (2007)
8. Popova, V., Sharpanskykh, A.: Modeling Organizational Performance Indicators. Information Systems 35(4), 505–527 (2010)
9. Pourshahid, A., Amyot, D., Chen, P., Weiss, M., Forster, A.J.: Business Process Monitoring and Alignment: An Approach Based on the User Requirements Notation and Business Intelligence Tools. In: WER, pp. 80–91 (2007)
10. ITU-T – International Telecommunications Union: Recommendation Z.150 (02/03), User Requirements Notation (URN) – Language Requirements and Framework, Geneva, Switzerland (2003)
11. Berander, P., Andrews, A.: Requirements Prioritization (Chapter 4). In: Aurum, A., Wohlin, C. (eds.) Engineering and Managing Software Requirements, pp. 69–94. Springer, Berlin (2005)
12. Lausen, S.: Software Requirements – Styles and Techniques. Pearson Education, Essex (2002)
13. Karlsson, J.,, Wohlin, C., Regnell, B.: An Evaluation of Methods for Prioritizing Software Requirements. Information and Software Technology 39(14-15), 939–947 (1998)
14. Brackett, J.W.: Software Requirements Technical Report. Carnegie Mellon University, Software Engineering Institute (1990)
15. Expert Business Analyst: Requirements Prioritization, `http://expertbusinessanalyst.com/requirements-prioritization/`
16. Karlsson, L., Berander, P., Regnell, B., Wohlin, C.: Requirements Prioritisation: An Experiment on Exhaustive Pair-Wise Comparisons vs. Planning Game Partitioning. In: Proceedings of the 8th International Conference on Empirical Assessment in Software Engineering (EASE 2004), pp. 145–154. IEE, Stevenage (2004)
17. Leffingwell, D., Widrig, D.: Managing Software Requirements – A Unified Approach. Addison-Wesley, Upper Saddle River (2000)
18. Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)
19. Regnell, B., Höst, M., Natt och Dag, J., Beremark, P., Hjelm, T.: An Industrial Case Study on Distributed Prioritisation in Market-Driven Requirements Engineering for Packaged Software. Requirements Engineering 6(1), 51–62 (2001)
20. Alshehri, S., Benedicenti, L.: Using the Analytical Hierarchy Process as a Ranking Tool for User Story Prioritization Techniques. In: Proceedings of the 8th International Conference on Software Engineering Advances, pp. 329–335. IARIA, Venice (2013)
21. A Guide to the Business Analysis Body of Knowledge® (BABOK® Guide), version 2. International Institute of Business Analysis, Toronto, Ontario, Canada (2009)

Combining Work Process Models to Identify Training Needs in the Prehospital Care Process

Eva Söderström[1], Joeri van Laere[1], Per Backlund[1], and Hanna Maurin Söderholm[2]

[1] School of Information Technology, University of Skövde, Box 408, 54128 Skövde, Sweden
[2] Swedish School of Library and Information Science, University College of Borås, 50190 Borås, Sweden
{eva.soderstrom,joeri.van.laere,per.backlund}@his.se, hanna.maurin@hb.se

Abstract. The prehospital process is complex and covers a wide range of locations, healthcare personnel, technologies and competences. Enabling high quality holistic training is hence a challenge. Process models are efficient tools for representing reality, but no single modeling approach can cover the complexity of prehospital care. In our research, we have investigated the possibility to combine various process modeling techniques in order to identify training components and as many perspectives of the prehospital process as possible. Results show that combining different approaches and adapting them based on the need at hand is a successful strategy for enabling an of the prehospital care process from multiple perspectives, including identification of holistic, realistic and engaging training components. Future work can utilize our results to build training scenarios that can be implemented in training using for example simulation.

Keywords: prehospital care process, process modeling, training components.

1 Introduction

Prehospital emergency care is an important link in the overall care chain, and one of the keys to providing reliable, efficient and good healthcare to citizens. This requires competent and efficient emergency medical personnel with relevant education and a broad range of skills that are up-to-date. There are, however, many challenges to effective and realistic training in this domain. Because of the complexity of the prehospital care process, current training approaches are not sufficient. In order to create holistic, realistic and engaging training, thorough knowledge and analysis of the prehospital care process is crucial. Very few studies have so far modeled the prehospital care process, with Jensen (2011) being an exception. In this paper we propose the use of different work process models for modeling this process, to enable the prehospital care process to be viewed from several perspectives.

B. Johansson, B. Andersson, and N. Holmberg (Eds.): BIR 2014, LNBIP 194, pp. 375–389, 2014.

1.1 The Prehospital Care Process

The term "prehospital care" covers a wide range of medical conditions, medical interventions, clinical providers and physical locations (IBTPHEM, 2012). Medical conditions range from minor illness and injury to life threatening emergencies. Prehospital interventions therefore also range from simple first aid to advanced emergency care and prehospital emergency anaesthesia (IBTPHEM, 2012).

The prehospital care process is complex, diverse and often fragmented in different ways, e.g. because it involves: moving and transporting patients between different locations, consists of (often) time-critical work; and, requires collaboration between actors from different professions and organizations. As discussed by Söderholm (2013), the overall chain is plagued by both redundancies, gaps and loss of information, as well as insufficient ICTs (e.g. Reddy, et al., 2008). Furthermore, Vessgren and Wahlberg (2010) have identified a number of specific challenges that prehospital personnel encounter on-scene:

- environmental factors, such as traffic, weather, time of day, sound levels, number of people present, handling patients on-scene, and security risks;
- lack of information, such as being insufficient or difficult to interpret information about for example number of injured, number of vehicles, and types of injuries;
- poor on-scene communication and prioritization of patients, due to problems with or infrequent use of prehospital care management on scene;
- communication difficulties: Remote communication can be problematic, such as when technical equipment does not work, or communication differences between different professions, but also on-scene between colleagues and with patients, as well as risk of missing information.

Hence, the prehospital work process is impacted by concrete factors (such as environmental), but many problems concern a lack of coordination, information, documentation, and general approaches. Failing to meet these challenges might have consequences further on in the care process. For example, lack of, or insufficient documentation by EMS personnel in the field has been associated with poor patient outcomes (Laudermilch et al, 2010). Also, seemingly brief and lightweight communication/information exchanges, such as when paramedics are dispatched to a call by SOS Alarm, or, when they call a nurse in the emergency room to pre-alert about arrival and basic patient categorization and status, are important for understanding what has happened in the field or during transport and indicate or even determine what will happen next. From a training perspective, it is crucial to look at the prehospital process as a whole, and thus including aspects not directly related to patient care procedures, such as transportation between different locations, communication and collaboration.

1.2 Training Needs in a Prehospital Context

Competence can be simply defined as the ability to operate to an adequate, safe standard (Clements & Mackenzie, 2014). Competence-based training and assessment has become central to education and training for healthcare professionals (Clements & Mackenzie, 2014). There is, however, some uncertainty about the concept of competence and how the principles underpinning competence based training and assessment can be applied to evolving subspecialty and multidisciplinary areas such as prehospital and retrieval medicine (Clements & Mackenzie, 2014).

The term “Prehospital Emergency Carer” can denote both registered nurses and paramedics working in prehospital emergency care (Sandman & Nordmark, 2006). Since 2005, the health authorities in Sweden (where this study takes place) have required at least one registered nurse in every prehospital emergency team. (Vessgren & Wahlberg, 2010). As discussed by Barley and Orr (1997) prehospital professions such as ambulance nurses and paramedics have undergone dramatic changes, from "just being the ambulance driver" to becoming professionals providing advanced prehospital patient care. Hence, the profession is located in-between the medical realm and a service occupation, in the sense that they are “the expert” when treating patients within their own mobile work setting: in the ambulance or at the location of an incident or patient’s home, and, at the same time, also “a servant” in the sense of transporting patients to higher levels of care, i.e. bringing patients to physicians and hospitals. (Barley & Orr, 1997). Hence, the prehospital work context requires different competences. Although focusing on emergency physicians, Chapman et al (2004) identified the following competence categories in prehospital care:

- Patient care competency: Timely, effective, appropriate, and compassionate
- Medical knowledge competency: To know and apply the sciences
- Practice-based learning competency: Analyze, assess and perform systematic practice experience
- Interpersonal and communication skills competency: Interaction with different roles and with different conditions
- Professionalism: Professional, ethical and sensitive to a diverse patient population.
- Systems-based practice competency: Awareness and use of system resources

The first step to enrich the training for prehospital work is to identify the components and dimensions that are relevant and necessary to include in realistic training. Training scenarios are currently very one dimensional, and fail to incorporate the richness of “real” prehospital cases.

1.3 Work Process Modeling Techniques

A process model is an abstraction of reality that attempts to represent its important aspects (Browning, 2010). Process models are used to represent processes and the ways of working, to gain a deeper understanding of processes and related data in order to for example obtain a comprehensive vision of the system, organization, possible changes and interventions (Söderström, 2001). Process models enable the identification of where problems arise; who should deal with them; and whose interests that would be hurt if problems are not managed (Yu & Mylopoulos, 1993). In this paper, we use process models to gain an in-depth understanding of the prehospital process, what activities that are included therein, the roles involved and the information related to the activities. Roles are commonly modeled in Swimlanes (vertical columns or rows), which is one way of ensuring that responsibilities for activities are included in the process models (Ferrante et al, 2013).

Alter's (2006; 2008) work system framework is a modeling approach to quickly grasp and demarcate the main purpose and elements of a work system. A work system consists of 6 central elements (customer, product, activities, participants, information and technology) and 3 supporting elements (infrastructure, strategy and environment). Alter developed this method to show information system designers that an information system consists of more than just information technology and that all 9 elements need to be addressed in analysis and design, not just technology (which is only one of nine elements). Alter (2006; 2008) clearly argues that the Work systems framework is suited to model any work system or activity, not only information systems. As such, it is a suitable approach for modeling the prehospital care process.

2 Research Method

2.1 Research Objective and Question

The setting of this paper is one large region in Sweden called the Västra Götaland Region (VGR). Region Västra Götaland (VGR) has identified a need to train the prehospital process in a way that better reflects its complexity with respect to both physical contexts and diversity in tasks. This need was expressed as a combination of medical treatment, transportation, communication and care taking throughout the entire chain of events from call-out to delivery at the emergency care unit (fig 1).

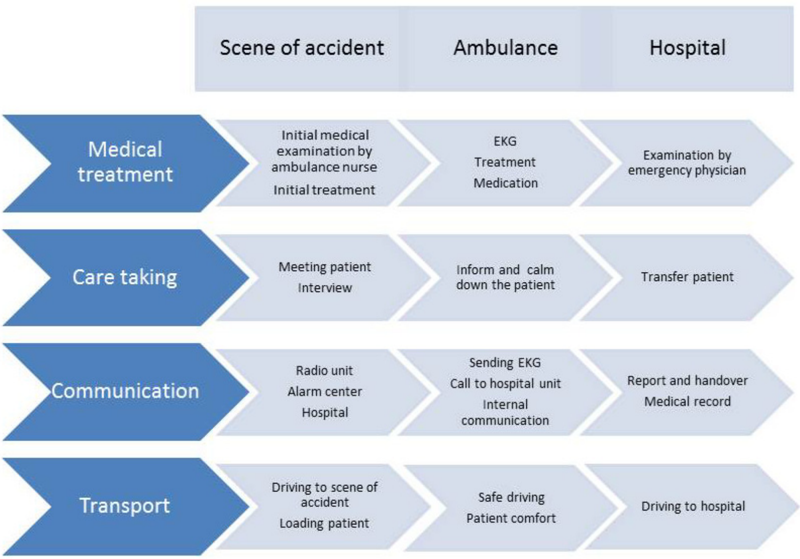


Fig. 1. Initial understanding of the prehospital care process

Figure 1 constitutes the initial understanding of the problem as created during a workshop where representatives from the prehospital care in VGR and simulation training researchers contributed. The figure served as a starting point for a pilot study carried out in 2012 where a training scenario and corresponding technology support were prototyped and tried out with the aim of exploring the feasibility of such an ambitious training scenario (Backlund et al, 2013).

After the successful pilot study, the next step was to develop a real demonstrator. A first step in developing this demonstrator was to study the prehospital care process in detail with the twofold aim of

1. describing the current work process as detailed as possible to be able to develop a realistic simulation environment that resembles the actual work situations, and
2. to identify training needs which need to be possible to pursue and evaluate in the simulation environment.

Given existing experience with Work Systems modeling (Alter, 2008) and process modeling (Ferrante et al, 2013) it was decided to visualize current work processes and training needs with these modeling techniques. Hence, our research question is "How can work systems and process modeling be used to describe and to identify training needs in the prehospital care process?" Process model combinations have previously proven successful, for example in resource modeling (Steele et al, 2001), and for creating process-aware information systems (Weidlich & Mendling, 2012). We argue that this is a beneficial approach for the purpose at hand.

2.2 Data Collection and Data Analysis

The construction of the models was an iterative process with a number of data collection activities (source triangulation) performed in parallel by three researchers with complementing background (expertise triangulation). Two researchers had a work process modeling background, where one is more experienced with Work System modeling and the other has more experience in Swimlane-modeling. The third researcher has a background in information sharing and ICT for prehospital care. Our work is focused on the models and what they can express, but elaboration of the way the models were developed is also important for validity and understandability.

The four models (see Chapter 3) were produced and refined in parallel and continuous comparison between them enabled us to enrich them a number of times. A literature review was done to identify existing prehospital care process models, which identified the model by Jensen (2011). This model gave a first rough understanding of important steps in the process and guided our interviews and observations.

Next, semi-structured interviews were performed with different domain experts. A 2-hour interview was conducted with two representatives for prehospital care training in VGR. The researchers spent six full days at an ambulance station interviewing ambulance personnel when they were not on emergency assignments. The interviews were geared to discover steps in the work process that ambulance personnel recognized, without guidance or constraints by any existing models.

In the first phase, the interviews were documented as running text. About halfway, the gathered data was structured in (1) a Work System model and (2) a Swimlane process model. Initially, these models were constructed independently of each other by two different researchers, and eventually they were compared and refined to become consistent. In the third phase of the interview period at the ambulance station "blind spots" in the models guided more in depth interviewing. The prehospital care researcher did not participate in the model construction, but in model evaluation and refinement by (1) questioning ambiguities in the models that conflicted with prior understanding or other contexts, and (2) posing complementing interview questions based on the more in-depth knowledge of the prehospital care process.

Three activities were conducted to further validate and refine the models: (1) participative observation during 10 emergency call assignments, during which the researchers discussed with the ambulance personnel, (2) one researcher spent one day at another ambulance station in Sweden, to check generalizability of our process, (3) presenting and using the models in various meetings: with the ambulance station's responsible doctor, and with researchers in training simulation.

The next section (3) shows how the models were used to visualize the current work process and to identify training needs. Elements were added to the work system framework to highlight conditions of the work environments and to accommodate for the special purpose of presenting training needs. We discuss these applications of the modeling methods further on in the paper.

3 Prehospital Care Work Process Models

The iterative modeling process resulted in four different ways of visualizing the prehospital care process:

1. A "general overview" process flow model at a high abstraction level (figure 2).
2. A detailed Swimlane process flow model (covering 10 printed A3 pages).
3. A Work Systems model of the complete process and 10 more detailed Work system models of the sub-phases in the overall process (examples in figure 3).
4. A Mapping model of the training needs identified in each of the 10 sub-phases in the Work system model across the four perspectives medical treatment, care taking, communication, transport (snapshot depicted in figure 4).

This chapter will present each of these in more detail, with an emphasis on their contents. In chapter 4, the analysis of the four models will be conducted, and the training needs identified elaborated upon.

3.1 General Overview Model

Very few studies have so far focused on modeling the prehospital care process. Jensen (2011), describes a process with a focus on decision support. We used this model as an inspiration, to allow for comparability between theory and practice. The Jensen (2011) model includes: Ambulance call dispatched, Paramedics en route, Paramedics arrival on scene, Patient contact, Assessment, On-scene treatment, Departure, En-route treatment, Arrival, and Patient hand-over. Based on the initial interviews with the education staff at the ambulance center, we developed a first version of a process model for the prehospital care process (Figure 2).

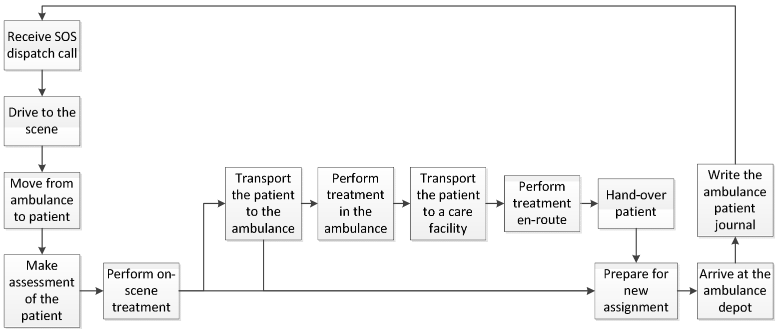


Fig. 2. General model of the prehospital process

The similarities between our model and the one by Jensen (2011) are clear, but some significant differences exist. Treatment of patients can occur in three places during the process: on-scene, in the ambulance on scene, and in the ambulance en-route. We have also added two activities at the end of the prehospital care process: *Write the ambulance patient journal*, and *Prepare for new assignment*. The journals are essential for follow-up and hand-over purposes, and are filled out during several activities in the process. The original document is handed over to the emergency care department, and a copy is kept and entered into the ambulance system once back at the ambulance depot. Data from SOS Alarm with time stamps are included and automatically inserted in the journal via systems integration. Besides writing the journals, the ambulance staff also needs to prepare for the next assignment. This includes both re-stocking of medical equipment, which can take place both at the emergency ward and at the ambulance depot, as well as cleaning the ambulance and discussing potentially stressful or difficult experiences from the case.

3.2 Detailed Swimlane Process Flow Model

The complexity of the prehospital care process was uncovered step by during the interview and observation process. Details about the different activities, decisions, documents, systems and information flows were discussed in-depth, and the result was documented in a detailed Swimlane process flow model. The Swimlanes are depicted as rows, and each role is assigned to one Swimlane. A role can be a human, a document or a system. The two persons who are the ambulance personnel are the key actors along with the patients. In addition, family, witnesses to accidents/events and the healthcare information organization can place calls for ambulances to SOS Alarm besides the patient him-/herself.

Various healthcare staff groups are or can be involved both during prehospital and in-hospital treatment and patient care. During the process, different IT systems are included, as well as different document templates and instructions. The Swimlane process flow model shows what activities that take place in the different roles/rows, and what information that flows in between, with an emphasis on the order in which things happen and what information that is used, produced, communicated and stored and where. The model itself ended up being the size of 10 A3 pages, and is hence too large to display in this paper. Mere snapshots cannot show sufficient detail to provide an understanding of the full complexity and how the model was used, and hence we made the decision to only describe this model in words.

3.3 Work Systems Model

In figure 3 an example of the Alter model for the sub-phase "Drive to the scene" is portrayed. Only the six most central Alter elements were modeled for each sub-phase. In addition two new elements were added (location and training needs) as will be discussed in section 4.2.

PRODUCT	CUSTOMER
Ambulance at address Status update from U (left station) to F (at address)	Ambulance personnel (medical + driver) SOS alarm (receives status updates) *Patient/relative (if they see the ambulance arrive)*

ACTIVITIES

- Ambulance-driver drives ambulance to address given route on navigation device
- If needed: Ambulance-medical-responsible assists with navigation
- Ambulance-driver and Ambulance-medical-responsible discuss strategy upcoming patient
- If needed: SOS Alarm updates with more detailed information on patient or address
- If needed: Ambulance-medical-responsible requests more details from SOS
- In case suspected risk for violence (known address): assist police assistance
- Ambulance-medical-responsible contacts patients in case of
  - No clear address: to obtain driving directions
  - Long distance: to get more information, give advice on treatment until ambulance arrives, relax patient, or give an estimate when the ambulance will arrive
- Ambulance personal confirms arrival at address: status changed to F (arrived)
- At arrival: windshield check (safe to go out?, patient visible?, other people?)

PARTICIPANT	INFORMATION	TECHNOLOGY
Ambulance-driver Ambulance-medical-responsible SOS alarm	• Address and route • More details address • More details where exactly patient is at address • More details on symptoms	• Computer with navigation device • Touchscreen • Radio communication tool

LOCATION (WHERE ARE WE – HOW IT LOOKS LIKE)	TRAINING IDEAS
• Ambulance personnel in front of ambulance • Siren noise – blue light reflection visible • At high speed unsteady • Computer between ambulance personnel's seats • Handbooks also between seats • Do we simulate 'their' area – they know common streets and city areas • At arrival there is a direct "windshield check" (how do we solve that in our simulator)	• Search for more information (SOS, handbooks, discussion with colleague) • Mentally prepared? What to do on arrival? • Difficulties to find address • How to prepare for "problematic" addresses • Communication discipline in radio • How to contact patient during drive

Fig. 3. Alter model for Drive to the scene

3.4 Mapping Model of Training Needs

The 10 sub-phases in the General overview process flow model was used as the basis in the mapping model. Based on Backlund et al (2013), the different physical locations were inserted above the process as input and as a contextual description affecting the design of prehospital training. Below each sub-phase, the four perspectives transport, communication, medical and "treatment" were inserted. Each sub-phase was analyzed using the four perspectives and if any relevant connections were identified, they were written down under these headings. Figure 4 shows part of the mapping model, where only the first three sub-phases are included for readability.

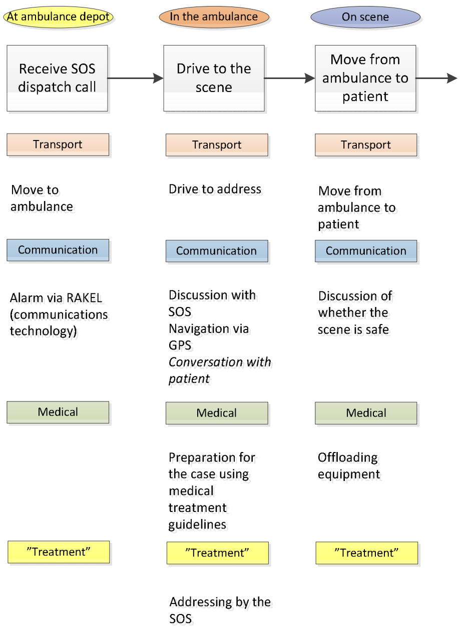


Fig. 4. Mapping model for training needs (snapshot)

4 Analysis of Training Components

4.1 Progression of the Process Models

The four models had different purposes and target groups. The general overview model should quickly present all included steps in one overview for the involved prehospital care domain experts, ambulance personnel or training simulation developers. One domain expert commented that 13 process steps still were too many for presentation at the higher decision making levels: "when we have to present this for the upper management we need a process with maximum 5 steps". This model was still the most effective for validating the correctness of the models, but insufficient for 1) understanding details in the prehospital process, 2) understanding roles and their interactions, 3) getting an overview of information flows, and 4) what technical systems and documents which are important.

The detailed Swimlane process flow model addresses these issues, but has shortcomings when it comes to 1) failure to account for different locations, and 2) failure to account for training needs. Locations set boundaries for accessibility to patients and equipment, and are associated with environmental factors that can affect the prehospital work. Location is hence an essential aspect in ambulance personnel training. The Alter models were the most useful ones to capture and show details of activities and all aspects that were not visualized in the process flow models. The Alter models were with respect to granularity between the overview model and the detailed Swimlane model, and as such most suited to discuss correctness and validity of our models with ambulance staff (where the overview model was too rough and the detailed Swimlane too large).

Thus far, the combination of a general overview model, a detailed Swimlane process flow model and a work systems model has enabled us to achieve a deep understanding of the prehospital process in terms of activities, roles, communication, information flow, documents, IT systems, and location. However, none of the three models show all the identified training needs at one glance and related to the four perspectives. Therefore the fourth model was customized for the meeting with prehospital care domain experts and training simulator developers, and also addressed the "human perspective" of the prehospital process, which in this case refers to how the ambulance personnel, patients, SOS personnel, care facility personnel etc address or "treat" each other. It concerns professional communication styles to be adopted and which need to be trained in order to be properly performed. This facet was highlighted by the VGR region (see figure 1) as an important factor. In order to cover this aspect, we went back to the general overview model and discussed it from the four perspectives described by the VGR, resulting in the mapping model. This model became the only model that shed light onto the prehospital process from these perspectives. Modeling the same process from different viewpoints and on different abstraction levels creates a deeper understanding of the process under study. Continuous comparison of the models resulted in refinement of all models. A lesson learned is hence that model combination is preferred before one way of modeling.

4.2 Modeling Challenges

An important modeling breakthrough was the creation of two extra elements (location and training needs) in the Work system model sheets. Neither the general overview process model, nor the Swimlane model and the original Work process model could visualize training needs. Normally, one would create two variants of each model, one "AS-IS" model and one "TO-BE" model. However, training needs do not always relate to a "change in the way of working", they can also relate to a particular part/step of the process experienced as "difficult". An alternative would have been to highlight these places in the models with a color, but the creation of a separate element was judged to be more helpful, as it also enabled to clarify why the training need was important. Another important aspect was "what the physical environment looked like". This was a very important type of information that needed to be captured for developing realistic and valid training, and it was not captured in either of the process models thus far. With some fantasy, the work system framework element "environment" could be used, but in Alter (2006) the "environment-element" is more referring to outside factors like "laws, technological trends in society" etc. Therefore, a specific customized element was created where "typicality of and constraints in the nearby physical environment" could be listed. Examples of issues listed here were "tight/narrow in the ambulance", "heavy to lift patient on the ambulance stretcher", "uncomfortable driving with stretcher on not plain ground outside a house" et cetera.

Several things were hard to model and required iteration to be solved. It took quite some time to decide whether the overall process should be divided in 8, 10 or 13 steps, e.g. which of these could/should be merged to one, and which needed to be separate. Sometimes the modeling approach decided which direction to go, e.g. in the Swimlane model where numerous interactions between people required splitting, whereas the same phase could not generate so much information in the work system elements (urging to merge two rather empty work system sheets in to one). In some case we actually chose to have different divisions in the different modeling approaches which did not lead to inconsistency, but rather a different level of detail. This is also visible in table 1, where the work system phases (10) not completely correspond to the 13 steps in the overall process overview.

Another difficult issue was to visualize the care taking as a separate activity. Care taking is often an integral part of other activities. For example, "the way you lift or carry a person on the stretcher" (relates to transport), "the way you perform medical examination" (relates to medical treatment), and "the way you communicate with the patient besides the necessary medical questions" (relates to communication). Abelsson et al (2014) surmise that prehospital personnels' relation to patients is often overlooked in research on prehospital simulation. Hence, although it is perfectly possible for an evaluator in training or in reality to judge care taking in general or in relation to specific activities, it is hard to highlight these issues in activity models, if not "goals" for these activities are included.

4.3 Identification of Training Components

Our approach can be used as a tool to identifying different training components – both as specific activities or isolated steps in the overall process, and their place within the overall process. This is done through the possibility to move between the different levels of detail in the different models. For example, the small model can be used to select an activity to train, then the large model can be used to zoom in and provide necessary details in order to create rich scenarios, and then the Work system models can be used to zoom out and describe what to train each step.

By combining the process models in our work, we have identified *what* to train in the prehospital process (medical skills, communication skills, treatment skills, and transportation skills), *when* to train them (i.e. during what activity in the process), and *where* to train them (in the ambulance depot, in the ambulance en-route, on scene, and at the care facility). The results show a richer, more holistic view of prehospital training that takes more aspects into account.

5 Concluding Discussion

This study has rendered results in three knowledge domains: an extended characterization of the prehospital care process; a concrete example of how process modeling can be used to derive training needs, and; and the benefits of process modeling combination. Each of these domains will be discussed in the following sub-sections.

5.1 Reflections Considering the Prehospital Care Process

Very few studies have so far focused on describing a process model for the prehospital care process. We used Jensen's (2011) model as a basis, in order to allow for comparability between theory and practice. The similarities and differences between our model and the one by Jensen (2011) are shown in table 1. The numbers reflect to the 10 work system sheets (which in some cases merge two activities that are shown as separate in the overall process flow model).

Table 1. Extending existing theory of the prehospital care process

Our study (2014)	**Jensen (2011)**
(1) Receive SOS dispatch call	Ambulance call dispatched
(2) Drive to the scene	Paramedics en route
(3) Move from ambulance to patient	Paramedics arrival on scene
	Patient contact
(4) Make assessment of the patient	Assessment
(5) Perform on-scene treatment	
(6) Transport the patient to the ambulance	On-scene treatment
(7) Perform treatment in the ambulance (before leaving the scene)	
(8) Transport the patient to care facility	Departure
(8) Perform on route treatment	En-route treatment
(9) Hand-over patient	Arrival
	Patient hand-over
(9) Prepare for new assignment	
(9) Drive to ambulance depot	
(10) Write the ambulance patient journal	

By incorporating different perspectives and different process modeling techniques, we were able to identify how the existing theory can be extended to a more holistic approach. The addition refers specifically to the added phases.

5.2 Reflections Considering Training Needs

As discussed, the prehospital work context requires a variety of different competences. One key challenge is the difficulty to explicate, represent and describe the many types of information people deal with in the prehospital care process. Taken together, the different types of information determine the paramedic's decisions on what course of action to take on interventions and patient care and where to transport the patient. (Söderholm, 2013). The complexity of the prehospital process and the different views thereof (figure 1) means that ambulance personnel need to be proficient in a number of areas, such as VGR's (Backlund et al, 2013) four perspectives: medical treatment, transportation, communication and care. This corresponds well with the competence needs identified by Chapman et al (2004), with the addition of transportation to their list.

Our key point is that it is the *combination* of these views is needed to identify training needs. Ambulance personnel must both know and apply medical treatments, and how to transport patients, themselves, equipment, information and the ambulance to and from various locations. They must be able to use systems and communication technology and enable communication in different ways, as well as be professional both when communicating with patients and with other healthcare actors. Therefore, future prehospital training must consider ways in which all perspectives can be taken into account to enable ambulance personnel to be better prepared.

The Alter models were used to identify training needs with an integrated view of the scenario in the form of a narrative and the technical components (i.e. the systems used by the staff, such as medical support systems and communication systems) involved. The quality of these narratives has been improved by the fact that they are directly derived from the actual process as creating these narratives would have been a less formalized process without them. Furthermore, the task of identifying technology support for the different phases of training was similarly improved by the fact that we had a clear view of which technologies (i.e. simulation and visualization technologies) where associated with what training goals. As an additional outcome we were also able to derive a number of technical requirements for many general features of a planned simulation and training facility.

5.3 Reflections Considering Process Modeling Combination

Process modeling is a very useful tool for describing and analyzing ways of working in organizations. The prehospital setting is not an exception, but rather benefits from process modeling in a variety of ways. In this paper, we have analyzed the complex prehospital care process by using different approaches to process modeling, with the aim of identifying training components for the ambulance personnel. Our combination of process modeling perspectives has not been identified in any previous research, and the results show the benefits of our approach for creating a holistic overview of what prehospital training for ambulance personnel must contain. Therefore, our results:

- can be used to identify and populate rich training scenarios that will provide ambulance personnel with a wider range of skills and more holistic training than currently available
- enrich current process modeling research by showing how process modeling variants can complement each other to capture the full complexity of processes, and that they can be used for different purposes and with different audiences
- extend in particular Alter's work processes with additional features that enable more information to be captured and hence also adds to their usefulness
- provide a tool to get a rich overview of the complex prehospital care process while, at the same time, allowing for in-depth focus on specific, isolated, aspects or steps and their place in the overall context.

All four types of process models are needed if the full potential for training in the prehospital process is to be realized. The general model can be used with upper management and politicians where an overview is sufficient to communicate the intended message. The larger detailed process model is needed to provide detail and navigation in the process, and allows training staff to select what part of the process to train and what they need to create good training in that part. The modified Alter descriptions serve a good purpose for discussions and details, while the mapping model provided different perspectives on the process model needed for training and stemming from the prehospital process competence needs.

5.4 Future Work

Future research can expand and elaborate on our results in a variety of ways: 1) Identifying rich training scenarios and elaborate on the possibility to use simulation in training, 2) Refine and describe in more detail what the different perspectives consist of, to for example separate communication via technology from human-to-human communication or professional language to other professions from professional language towards patients, 3) Elaborate on the use of additional process modeling techniques such as BPMN, and 4) Further analyze the expansion to the Alter work processes and use additional empirical studies to verify and refine our findings.

References

Abelsson, A., Rystedt, I., Suserud, B.O., Lindwall, L.: Mapping the use of simulation in prehospital care-a literature review. Scand. J. Trauma. Resusc. Emerg. Med., 22 (2014)

Alter, S.: The Work System Method: Connecting People, Processes, and IT for Business Results. Work System Press, Larkspur (2006)

Alter, S.: Defining information systems as work systems: implications for the IS field. EJIS 17, 448–469 (2008)

Backlund, P., Engström, H., Johannesson, M., et al.: Ambulansträningscenter. Förstudie prehospitalt tränings- och simuleringscenter för Västra Götaland. Technicalreport HS-IKI-TR-13-001 (2013)

Barley, S.R., Orr, J.J.E.: Between Craft and Science: Technical Work in U.S. Settings. Cornell University Press, Ithaca (1997)

Browning, T.: On the alignment of purposes and views of process models in project management. J. Oper. Manag. 28, 316–332 (2010)

Chapman, D., Hayden, S., Sanders, A., Binder, L., Chinnis, L., Corrigan, K., LaDuca, T., Dyne, P., Perina, D., Smith-Coggins, R., Sulton, L., Swing, S.: Integrating the accreditation council for graduate medical education core competencies into the model of the clinical practice of emergency medicine. Ann. Emerg. Med. 43, 756–769 (2004)

Clements, R., Mackenzie, R.: Competence in prehospital care: evolving concepts. Emerg. Med. J. 22, 516–519 (2005)

Ferrante, S., Bonacina, S., Pinciroli, F.: Modeling stroke rehabilitation processes using the Unified Modeling Language (UML). Comput. Biol. Med. 43, 1390–1401 (2013)

IBTPHEM (Intercollegiate Board for Training in Pre-hospital Emergency Medicine): Sub-specialty training in pre-hospital emergency medicine: A guide for trainees, trainers, local education providers, employers and deaneries. IBTPHEM, London (2012)

Jensen, J.: Paramedic clinical decision-making: results of two Canadian studies. Int. J. Param. Pract., 186–194 (2011)

Laudermilch, D.J., Schiff, M.A., Nathens, A.B., Rosengart, M.R.: Lack of emergency medical services documentation is associated with poor patient outcomes: a validation of audit filters for prehospital trauma care. J. Am. Coll. Surgeons 210(2), 220–227 (2010)

Reddy, M.C., Paul, S.A., Abraham, J., McNeese, M., DeFlitch, C., Yen, J.: Challenges to effective crisis management: Using information and communication technologies to coordinate emergency medical services and emergency department teams. Int. J. Med. Inform. 78(4), 259–269 (2008)

Sandman, L., Nordmark, A.: Ethical conflicts in prehospital emergency care. Nurs. Eth. 13(6), 592–607 (2006)

Steele, J., Son, Y.-J., Wysk, R.: Resource modelling for the integration of the manufacturing enterprise. J. of Manufacturing Systems 19(6), 407–427 (2001)

Söderholm, H.M.: Emergency visualized: exploring visual technology for paramedic-physician collaboration in emergency care. Valfrid, Borås (2013)

Söderström, E.: Implications of Standardisation Efforts for Business Process Modeling. In: Kangassalo, H., Welzer, T., Jaakkola, H., Rozman, I., Kawaguchi, E. (eds.) Fr. Art. Int., pp. 105–112. IOS Press, Amsterdam (2002)

Vessgren, A., Wahlberg, M.: The ambulance nurse's experiences of problems on scene. BSc Thesis, Uppsala University, Uppsala (2010) (in Swedish)

Weidlich, M., Mendling, J.: Perceived consistency between process models. Inf. Sys. 37(2012), 80–98 (2012)

Yu, E., Mylopoulos, J.: An Actor Dependency Model of Organizational Work - With Application to Business Process Reengineering. In: Kaplan, S. (ed.) Proceedings of the Conference on Organizational Computing Systems (COOCS 1993), Milpitas, pp. 258–268 (1993)

Author Index