

Design of Business Process Model Repositories

Requirements, Semantic Annotation Model and Relationship Meta-model

Mturi Elias

Academic dissertation for the Degree of Doctor of Philosophy in Computer and Systems Sciences at Stockholm University to be publicly defended on Monday 15 June 2015 at 13.00 in Lilla hörsalen, NOD-huset, Borgarfjordsgatan 12.

Abstract

Business process management is fast becoming one of the most important approaches for designing contemporary organizations and information systems. A critical component of business process management is business process modelling. It is widely accepted that modelling of business processes from scratch is a complex, time-consuming and error-prone task. However the efforts made to model these processes are seldom reused beyond their original purpose. Reuse of business process models has the potential to overcome the challenges of modelling business processes from scratch. Process model repositories, properly populated, are certainly a step toward supporting reuse of process models.

This thesis starts with the observation that the existing process model repositories for supporting process model reuse suffer from several shortcomings that affect their usability in practice. Firstly, most of the existing repositories are proprietary, therefore they can only be enhanced or extended with new models by the owners of the repositories. Secondly, it is difficult to locate and retrieve relevant process models from a large collection. Thirdly, process models are not goal related, thereby making it difficult to gain an understanding of the business goals that are realized by a certain model. Finally, process model repositories lack a clear mechanism to identify and define the relationship between business processes and as a result it is difficult to identify related processes.

Following a design science research paradigm, this thesis proposes an open and language-independent process model repository with an efficient retrieval system to support process model reuse. The proposed repository is grounded on four original and interrelated contributions: (1) a set of requirements that a process model repository should possess to increase the probability of process model reuse; (2) a context-based process semantic annotation model for semantically annotating process models to facilitate effective retrieval of process models; (3) a business process relationship meta-model for identifying and defining the relationship of process models in the repository; and (4) architecture of a process model repository for process model reuse.

The models and architecture produced in this thesis were evaluated to test their utility, quality and efficacy. The semantic annotation model was evaluated through two empirical studies using controlled experiments. The conclusion drawn from the two studies is that the annotation model improves searching, navigation and understanding of process models. The process relationship meta-model was evaluated using an informed argument to determine the extent to which it meets the established requirements. The results of the analysis revealed that the meta-model meets the established requirements. Also the analysis of the architecture against the requirements indicates that the architecture meets the established requirements.

Keywords: *business process, business process model, business process repository, semantic annotation.*

Stockholm 2015
<http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-117035>

ISBN 978-91-7649-192-8
ISSN 1101-8526



Stockholm
University

Department of Computer and Systems Sciences

Stockholm University, 164 40 Kista

DSV Report Series

Design of Business Process Model Repositories:

Requirements, Semantic Annotation Model and Relationship

Meta-model

No. 15-011



Design of Business Process Model Repositories: Requirements, Semantic Annotation Model and Relationship Meta-model

Mturi Elias

©Mturi Elias, Stockholm 2015
REPORT SERIES NO. 15-011
ISSN 1101-8526
ISBN 978-91-7649-192-8

Printed in Sweden by Holmbergs, Malmö 2015
Distributor: Department of Computer and Systems Sciences

dedicated to my wife, Selestina
to my daughters, Glory, Grace and Gladness

Abstract

Business process management is fast becoming one of the most important approaches for designing contemporary organizations and information systems. A critical component of business process management is business process modelling. It is widely accepted that modelling of business processes from scratch is a complex, time-consuming and error-prone task. However the efforts made to model these processes are seldom reused beyond their original purpose. Reuse of business process models has the potential to overcome the challenges of modelling business processes from scratch. Process model repositories, properly populated, are certainly a step toward supporting reuse of process models.

This thesis starts with the observation that the existing process model repositories for supporting process model reuse suffer from several shortcomings that affect their usability in practice. Firstly, most of the existing repositories are proprietary, therefore they can only be enhanced or extended with new models by the owners of the repositories. Secondly, it is difficult to locate and retrieve relevant process models from a large collection. Thirdly, process models are not goal related, thereby making it difficult to gain an understanding of the business goals that are realized by a certain model. Finally, process model repositories lack a clear mechanism to identify and define the relationship between business processes and as a result it is difficult to identify related processes.

Following a design science research paradigm, this thesis proposes an open and language-independent process model repository with an efficient retrieval system to support process model reuse. The proposed repository is grounded on four original and interrelated contributions: (1) a set of requirements that a process model repository should possess to increase the probability of process model reuse; (2) a context-based process semantic annotation model for semantically annotating process models to facilitate effective retrieval of process models; (3) a business process relationship meta-model for identifying and defining the relationship of process models in the repository; and (4) architecture of a process model repository for process model reuse.

The models and architecture produced in this thesis were evaluated to test their utility, quality and efficacy. The semantic annotation model was evaluated through two empirical studies using controlled experiments. The conclusion drawn from the two studies is that the annotation model improves searching, navigation and understanding of process models. The process relationship meta-model was evaluated using an informed argument to determine the extent to which it meets the established requirements. The results of the analysis revealed that the meta-model meets the established requirements. Also the analysis of the architecture against the requirements indicates that the architecture meets the established requirements.

Abstrakt

Processhantering, också kallat ärendehantering, har blivit en av de viktigaste ansatserna för att utforma dagens organisationer och informationssystem. En central komponent i processhantering är processmodellering. Det är allmänt känt att modellering av processer kan vara en komplex, tidskrävande och felbenägen uppgift. Och de insatser som görs för att modellera processer kan sällan användas bortom processernas ursprungliga syfte. Återanvändning av processmodeller skulle kunna övervinna många av de utmaningar som finns med att modellera processer. En katalog över processmodeller är ett steg mot att stödja återanvändning av processmodeller.

Denna avhandling börjar med observationen att befintliga processmodellkataloger för att stödja återanvändning av processmodeller lider av flera brister som påverkar deras användbarhet i praktiken. För det första är de flesta processmodellkatalogerna proprietära, och därför kan endast katalogägarna förbättra eller utöka dem med nya modeller. För det andra är det svårt att finna och hämta relevanta processmodeller från en stor katalog. För det tredje är processmodeller inte målrelaterade, vilket gör det svårt att få en förståelse för de affärs mål som realiserats av en viss modell. Slutligen så saknar processmodellkataloger ofta en tydlig mekanism för att identifiera och definiera förhållandet mellan processer, och därför är det svårt att identifiera relaterade processer.

Utifrån ett designvetenskapligt forskningsparadigm så föreslår denna avhandling en öppen och språkoberoende processmodellkatalog med ett effektivt söksystem för att stödja återanvändning av processmodeller. Den föreslagna katalogen bygger på fyra originella och inbördes relaterade bidrag: (1) en uppsättning krav som en processmodellkatalog bejöver uppfylla för att öka möjligheterna till återanvändning av processmodeller; (2) en kontextbaserad semantisk processannoteringsmodell för semantisk annotering av processmodeller för att underlätta effektivt återvinnande av processmodeller; (3) en metamodel för processrelationer för att identifiera och definiera förhållandet mellan processmodeller i katalogen; och (4) en arkitektur av en processmodellkatalog för återanvändning av processmodeller. De modeller och den arkitektur som tagits fram i denna avhandling har utvärderats för att testa deras användbarhet, kvalitet och effektivitet. Den semantiska annotationsmodellen utvärderades genom två empiriska studier med kontrollerade experiment. Slutsatsen av de två studierna är att modellen förbättrar sökning, navigering och förståelse för processmodeller. Metamodellen för processrelationer utvärderades med hjälp av ett informativt argument för att avgöra i vilken utsträckning den uppfyllde de ställda kraven. Resultaten av analysen visade att metamodellen uppfyllde dessa krav. Även analysen av arkitekturen indikerade att denna uppfyllde de fastställda kraven.

Acknowledgements

Thank you God for everything!

Firstly, I would like to express my sincere gratitude to my main supervisor, Dr. Ilia Bider, for guidance and motivation over years. Secondly, my gratitude goes to my second supervisor, Professor Paul Johannesson, for his valuable advice and encouragement. I am grateful that Ilia and Paul became my supervisors and gave me guidance on conducting my research work. I thank Paul for his constant encouragement and precise advice all the way through. He helped me enormously in clarifying research issues, especially during the initial phases of my research work.

Thanks to the Swedish International Development Agency (SIDA) for funding this thesis project. My appreciation goes to the University of Dar es Salaam for granting me this opportunity. Also, I would like to thank all the personnel at Department of Computer and Systems Sciences (DSV) for providing a pleasant environment in which to conduct my research activities.

I am lucky to have had knowledgeable, friendly and helpful colleagues during my work at Information Systems Laboratory (SYSLAB) and DSV in general. I especially acknowledge Khurram Shazhad for our productive collaboration and for inspiring and constructive discussions during this research project.

I must acknowledge the intern student at the Invention and Technological Ideas Development Organization (ITIDO) who worked very hard to develop a prototype of the annotation tool used to evaluate the semantic annotation model. Thank you, Ezra. I also thank Dr. Charles Sokile for proofreading the thesis.

I would like to thank my parents for their love and support.

In closing, I would like to thank, in a special way, my wife Selestina for accepting and allowing me to be away from the family for several years. Thanks a lot for looking after our kids. In the same vein, I would like to thank our daughters, Gloria, Grace and Gladness, for accepting a roving dad at such a tender age when you needed him most! Gladness bore the pain of coming into the world when Dad was far away!

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Problem Statement	3
1.3	Research Goals and Approach	3
1.4	Summary of the Contributions	4
1.5	Publications	6
1.6	Structure of the Thesis	9
1.7	Conceptual Framework Used in the Thesis	12
2	Theoretical and Technical Foundations	16
2.1	Overview	16
2.2	Business Process Management	16
2.2.1	Business Processes	16
2.2.2	BPM Foundation	19
2.2.3	The BPM Life Cycle	21
2.3	Business Process Modelling	23
2.3.1	Process Representation Perspectives	25
2.3.2	Modelling and the Need for Reuse	27
2.4	Business Process Model Repository	29
2.4.1	Process Knowledge and Adaptability	30
2.4.2	Process Representation and Understanding	34
2.4.3	Repository Structuring and Retrieval	37
2.5	Summary and Discussion	38
3	Research Design	40
3.1	Research Paradigm	40
3.1.1	Design Science	42
3.2	The Research Process	44
3.2.1	The Research Process for the Overall Research Goal	44
3.2.2	The Research Process for Goal 2	49
3.2.3	The Research Process for Goal 3	50
4	Requirements for a Process Model Repository	54
4.1	Overview	54
4.2	Requirements Elicitation from Stakeholders	54

4.2.1	Exploratory Study	56
4.2.2	Confirmatory Study	59
4.3	Requirements Elicitation from Literature Survey	64
4.3.1	Planning	64
4.3.2	Conduction	65
4.3.3	Requirement Establishment.....	66
4.4	Requirements Specification and Justifications	67
4.5	Related Work.....	71
4.6	Summary and Discussion	72
5	A Context-based Process Semantic Annotation Model.....	73
5.1	Overview.....	73
5.2	Requirements for the Annotation Model.....	73
5.3	Annotation Model Development.....	75
5.3.1	Identification of Potential Annotation Elements	75
5.3.2	Validation of Potential Annotation Elements	76
5.3.3	Model Construction.....	77
5.4	CPSAM – Context-Based Process Semantic Annotation Model	78
5.4.1	Functional Perspective	79
5.4.2	Organizational Perspective.....	82
5.4.3	Informational Perspective	83
5.4.4	Business Process Context Perspective	84
5.5	Demonstration of the CPSAM Application.....	85
5.6	Related Work.....	92
5.7	Summary and Discussion	93
6	A Prototype of the Semantic Annotation Tool	94
6.1	Overview.....	94
6.2	Functionalities of the System	94
6.3	Activiti	95
6.4	The Prototype System	95
6.4.1	Process Modelling Environment	96
6.4.2	Semantic Annotation Service.....	99
6.4.3	Advanced Process Retrieval.....	99
6.5	Summary and Discussion	103
7	Evaluation of the Semantic Annotation Model.....	104
7.1	Overview.....	104
7.2	The Evaluation Framework.....	104
7.3	Study I – Validation of Annotation Correctness and Consistency.....	106
7.3.1	Experimental Settings.....	106
7.3.2	Studied Variables	108
7.3.3	Results and Discussion	109

7.4	Study II – Validation of User Performance and Perception	111
7.4.1	Experimental Settings.....	111
7.4.2	Studied Variables	113
7.4.3	Results and Discussion	116
7.5	The Evaluation.....	122
7.6	Summary and Discussion	123
8	Business Process Relationship: The Meta-model.....	124
8.1	Overview.....	124
8.2	Requirements for the Process Relationship Meta-model	124
8.3	A Process Relationship Meta-model Development	125
8.4	The Process-Assets and Asset-Processes Archetypes	128
8.4.1	The Process-Assets Archetype for Main Processes	128
8.4.2	The Asset-Processes Archetype	131
8.4.3	Archetypes for Supporting Processes.....	132
8.5	Validation of the Archetypes	133
8.5.1	Study Settings	133
8.5.2	Archetypes Validation Process	134
8.5.3	Educational Process Architecture: Results	136
8.5.4	Analysis of the Results	138
8.6	Business Process Relationship: Meta-model	142
8.7	Evaluation of the Meta-model	145
8.8	Summary and Discussion	147
9	Architecture for the Process Model Repository	148
9.1	Overview.....	148
9.2	Architecture Design and Development.....	148
9.2.1	Technical Problem Analysis.....	149
9.2.2	Solution Domain Analysis	153
9.2.3	Creating the Repository Architecture	157
9.2.4	The Repository Architecture	158
9.3	Architecture Specification	159
9.3.1	Data and Information Viewpoint.....	159
9.3.2	Functional Viewpoint	160
9.3.3	Standard Viewpoint	164
9.3.4	Workflow Activities.....	165
9.4	Architecture Evaluation.....	168
9.5	Related Work.....	169
9.6	Summary and Discussion	170
10	Conclusion and Future Work	172
10.1	Research Goals and Findings.....	172
10.2	Research Contributions	174

10.2.1	Contribution 1: Requirements for a business process model repository .	175
10.2.2	Contribution 2: A context-based process semantic annotation model.....	175
10.2.3	Contribution 3: A business process relationship meta-model.....	176
10.2.4	Contribution 4: Architecture for a business process model repository.....	177
10.3	Limitations	178
10.4	Future Research Directions	179
Appendices		180
A.1.	Survey of Existing Process Model Repositories	180
A.2.	Requirements Elicitation Questionnaire.....	188
A.3.	Requirements Confirmatory Study Instruments	190
A.4.	Validation of Potential Annotation Elements	193
A.5.	CPSAM Evaluation Instruments	195
A.6.	Case Study Interview Guide	197
References.....		200

Table of Figures

Figure 1: Structure of the thesis.....	10
Figure 2: A conceptual framework of the key concepts used and their relationships.	13
Figure 3: BPM life cycle, taken from (Dumas et al, 2013).	22
Figure 4: An order fulfilment process model example.	24
Figure 5: Ship and invoice subprocess of the order fulfilment process	24
Figure 6: Process model reuse (adapted from software reuse (Constantopoulos et al., 1995)).....	29
Figure 7: Process model fragment for invoice checking.....	32
Figure 8: Process model fragment for reuse of process logic (Schumm et al, 2011a).	33
Figure 9: A design science research process (Peffer et al, 2007).	41
Figure 10: A template of design science research process (Joahannesson & Perjons, 2012).	43
Figure 11: Research process for the overall research goal.....	47
Figure 12: Research process for achieving goal 2 of the thesis.....	48
Figure 13: Research process for achieving goal 3 of the thesis.....	53
Figure 14: Stakeholders' requirement elicitation process.....	55
Figure 15: Frequency distributions of propositions (P1–P12).....	60
Figure 16: Researchers and practitioners comparison on proposition P3.....	62
Figure 17: Researchers and practitioners comparison on proposition P5.....	62
Figure 18: Requirement elicitation – a systematic review process.....	64
Figure 19: Annotation model development process.	75
Figure 20: Validation of Concepts.....	77
Figure 21: A context-based process semantic annotation model (CPSAM). .	79
Figure 22: Order-to-cash business processes.....	86

Figure 23: Annotation of process model shown in Figure 22 (a).	88
Figure 24: Search result of query 1.	90
Figure 25: Search result of query 2.	91
Figure 26: Architecture of a prototype system.	96
Figure 27: Screenshot for creating and storing a new process model.	97
Figure 28: Screenshot for annotating a process model.	98
Figure 29: Searching a process model from the repository.	101
Figure 30: Process annotation of process model.	102
Figure 31: The Method Evaluation Model (MEM) (Moody, 2003).	105
Figure 32: Annotation consistency and correctness.	110
Figure 33: Perceived ease of annotation.	111
Figure 34: Search correctness.	117
Figure 35: Navigation efficiency.	118
Figure 36: Process model understandability.	120
Figure 37: User perception of the annotation model.	120
Figure 38: Annotation discriminating power.	121
Figure 39: The process-assets archetype for main processes.	130
Figure 40: An instantiation of process-assets archetype for teaching and learning as the main process.	130
Figure 41: An instantiation of the asset-processes archetypes for "student" asset.	131
Figure 42: An instantiation of the process-assets archetype for acquiring students.	132
Figure 43: Archetypes validation process.	134
Figure 44: An instantiation of process-assets and asset-processes archetypes.	140
Figure 45: A business process relationship meta-model.	143

Figure 46: The Architecture Synthesis Model (Tekinerdogan & Aksit, 2002).
..... 149

Figure 47: Repository system flow chart..... 157

Figure 48: Architecture of the process model repository. 158

Figure 49: Process information mediator..... 163

Figure 50: Populating the repository by creating a process model..... 165

Figure 51: Search process model activity diagram..... 166

Figure 52: Accessing process models from external repositories. 167

List of Tables

Table 1: Business process definitions drawn from the literature	17
Table 2: The study participants	56
Table 3: The derived propositions.....	58
Table 4: Propositions analysis (P1–P12).....	61
Table 5: Analysis of variance between practitioners and researchers	62
Table 6: Requirements (from stakeholders)	63
Table 7: Selected primary studies.....	65
Table 8: Requirements (from the literature).....	66
Table 9: Items for measuring perception-based variables.....	109
Table 10: Annotation consistency	110
Table 11: Annotation correctness	110
Table 12: Items for measuring perception-based variables.....	115
Table 13: Search correctness	116
Table 14: Navigation Efficiency.....	118
Table 15: Process model understandability.....	119
Table 16: Generalized requirements (technical problems).....	151
Table 17: The solution domain for the subproblems.....	153
Table 18: A selected set of the identified knowledge sources of the solution domain.....	154
Table 19: A set of knowledge sources for solution domain business process model abstraction technique (SD2).....	155
Table 20: Repositories review summary.....	186

Abbreviations

AD	Architectural Description
BPEL	Business Process Execution Language
BPM	Business Process Management
BPML	Business Process Modelling Language
BPMM	Business Process Metadata Model
BPMN	Business Process Model and Notation
BPMO	Business Process Management Ontology
BPMS	Business Process Management System
CPSAM	Context-based Process Semantic Annotation Model
EPC	Event-Driven Process Chain
ERP	Enterprise Resource Planning
IEEE	Institute of Electrical and Electronics Engineers
ISO	International Organization for Standardization
OWL	Web Ontology Language
PAIS	Process-Aware Information Systems
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
sBPEL	Semantically enhanced Business Process Execution Language
SBPM	Semantic Business Process Management
sBPMN	Semantically enhanced Business Process Modelling Notation
SBPR	Semantic Business Process Repository
SCM	Supply Chain Management
SCOR	Supply Chain Operations Reference
sEPC	Semantically enhanced Event-Driven Process Chain
SOA	Service-Oriented Architecture
UML AD	Unified Modelling Language Activity Diagram
UN/CEFACT	United Nations Centre for Trade Facilitation and Electronic Business
WFM	Workflow Management
WFMS	Workflow Management System
YAWL	Yet Another Workflow Language

1 Introduction

Business process management (BPM) is increasingly being adopted as an approach for designing contemporary organizations and information systems (Aldin & de Cesare, 2011). Business process modelling is a primary requirement and one of the critical components for successful BPM implementation. While the modelling of business processes remains a complex, costly and time-consuming task (Rodrigues et al, 2006; Markovic & Pereira, 2008b; Hornung et al, 2009), efforts made to model business processes are seldom reused beyond their original purpose. An attractive approach to modelling business processes from scratch is deriving them by redesigning existing models. Process model repositories, properly populated, are certainly a step towards process model reuse.

This thesis addresses the design of a process model repository as an infrastructure for storing, managing and sharing of process models for future reuse. In this chapter, we motivate the research topic and explain the research problem. Furthermore, a research goal is specified together with the contributions of the research.

1.1 Motivation

Over the last decade, due to the rapid development of Internet technologies, enterprises have extended their functions to customers, business partners and financial institutions (Zdravkovic, 2006). As a result, enterprises are relying on more complex systems than ever. This requires enterprises to continually streamline and align their business processes with partner processes. Also, the information technology activities and infrastructures need to be centred around the business processes in order to attain a better business performance – maximum efficiency, short lead times, etc. Therefore, the interest in workflow systems and business process management (BPM) has been steadily increasing.

In BPM, the management of business activities is based on a framework of operational processes. A business process is a chain of activities performed in an organization that ultimately add value for its customers (Weske, 2007). Business process modelling is an approach to explicitly represent the way organizations conduct their business operations (Indulska et al, 2009). As pointed out by Indulska et al (2009), business process modelling is a primary requirement for organizations desiring to adopt BPM. A process model, typically given in graphical form, describes the activities,

events and control flow logic that constitute a business process (Recker et al, 2009). The models may include additional information such as business goals, business performance metrics etc. In addition, process models are the basis for the analysis and design of Process-Aware Information Systems (PAISs). A PAIS is a “software system that manages and executes operational processes involving people, applications, and/or information sources on the basis of process models” (Dumas et al, 2005).

Abstracting and making the process logically explicit through process models offers several benefits, including (Dumas et al, 2005; van Der Aalst & Van Hee, 2004; Indulska et al, 2009):

- *Maintained focus on business needs.* During information systems analysis and design, the focus is kept on the business processes and not their technical realizations. This focus promotes clear communication with business users and facilitates the alignment of information systems to business requirements.
- *Automated enactment.* The explicit representation of the business processes through process models enables their automated enactment in software, which leads to improved business performance by rationalizing the use of the available resources.
- *Easy change management.* When a business process changes, reflecting these changes in the model will trigger the alignment of the underlying systems.
- *Process improvement.* Process models provide organizations with a greater ability to understand and improve their business processes.
- *Management support.* Process models enable management support at the design and control levels through simulation, monitoring and process mining facilities.

While the modelling of business processes provides enormous benefits, it is a complex, time-consuming and error-prone task (La Rosa et al, 2011a; Rodrigues et al, 2006; Markovic & Pereira, 2008b). The reasons for this may include: (a) the high intrinsic complexity of many business processes and (b) the difficulty of reaching agreement on how the processes will be run between multiple stakeholders, with differing interests and goals, involved in the design. While it may be difficult to address the second issue, we believe that there are effective solutions for managing the complexity of business processes. One possible solution is to collect and share process models through the use of a process model repository. Such a repository may contain process models and their associated process knowledge, each of which may describe a specific business process. The main benefits of such a repository include process model reuse and knowledge exchange.

1.2 Problem Statement

Process model reuse via repositories is not a new concept. There has been several efforts to build such repositories. Some of the widely known repositories include the MIT process handbook (MIT process handbook project, 2001), SAP's Business Map (SAP AG, 2007) and the IBM process repository (IBM Corporation, 2004). Despite their benefits, the existing process model repositories suffer from several shortcomings that affect their usability in practice. During the initial phase of this research we carried out a survey (see Appendix A.1) to identify some of the main challenges that limited existing process model repositories from supporting process model reuse. The results of the survey indicate that lack of an effective retrieval system is common to all repositories. Also, some of the recent studies (Yan et al, 2010; Yan, 2012) have indicated that the lack of an efficient retrieval system is one of the main problems affecting existing repositories. In addition, the survey has indicated that most of the existing repositories are proprietary – they are not publicly open. This means they do not allow users to add new models or modify existing ones without any prior legal permission, which leads to a limitation that only the owners of the repositories can extend or enhance them with new process models. Consequently we lack a critical mass of process models that are available for reuse. Furthermore, we lack transparency between different repositories, which could allow shared use as a standard resource.

We envision that the above-mentioned limitations can be overcome through an open and language-independent repository with an efficient retrieval system – a repository that is publicly open to any potential user, independently of modelling language, and can comprise process models from existing process repositories. Therefore the overall problem this thesis addresses is *how to design an open and language-independent process model repository with an efficient retrieval system to support reuse of process models.*

1.3 Research Goals and Approach

The main goal of this research is “to design an open and language-independent process model repository with an efficient retrieval system to support reuse of process models”. More specifically, the main goal is decomposed into the following four goals.

- Goal 1. To establish the requirements of a process model repository for process model reuse
- Goal 2. To develop and evaluate a process semantic annotation model for semantically annotating process models in the repository

- Goal 3. To develop and evaluate a model for identifying and defining relationships between process models in the repository
- Goal 4. To design the architecture of a publicly open and language-independent process model repository

To achieve the above research goals, we have followed a design science research paradigm (Hevner et al, 2004). This is because the goal of design science is to create information technology artefacts to solve organizational problems (Hevner et al, 2004). This is in line with our research goals, which aimed at developing three artefacts to address some of the challenges affecting in existing repositories. The created artefacts includes a semantic annotation model for semantically annotating process models in the repository to facilitate efficient retrieval of process models; a business process relationship meta-model for identifying and defining relationships between process models; and the architecture of the process model repository.

The construction of an artefact does not directly provide any empirical data, however the knowledge that is required in the design process comes from empirical and theoretical sources. Therefore, several complementary research strategies and methods have been employed to prepare and evaluate the design process and to collect the necessary data. In addition, during problem identification additional research strategies and methods were needed to gain a sufficient empirical basis. The complementary research strategies employed in this thesis include survey, case study and laboratory experiments (Hevner et al, 2004; Peffers et al, 2012), whereas the research methods employed include interview, questionnaire and document analysis.

1.4 Summary of the Contributions

Following the deployed research paradigm, strategies and methods, this research has resulted in the following major contributions:

1. A set of requirements that must be fulfilled by process model repositories in order to increase the probability of process model reuse.

One of the contributions of this thesis is a set of requirements that must be fulfilled by process model repositories in order to increase the probability of process model reuse. The purpose of these requirements is to guide the design and development of the process model repository. In addition, the established requirements for a process model repository provides a better understanding of the problems that affect existing process model repositories in supporting reuse. The requirements were elicited from stakeholders and the literature through a systematic review approach. While some definitions of requirements for process model repositories existed, the elicitation of such

requirements from a group of stakeholders is new. In addition, since process repositories can be designed for several different purposes, requirements for the repository to support reuse of process models are new. Therefore the set of requirements definitions provided in this thesis serves as an extension and validation of existing definitions of the requirements. It should be noted that the presented requirements could be extended and adapted based on the primary purpose of the repository.

2. A context-based process semantic annotation model (CPSAM) for annotating process models in the repository.

As the main contribution of this research, the semantic annotation model is developed to semantically annotate process models in the repository. The purpose of the annotation model is to facilitate searching for process models, navigating the repository and enhancing user understanding of process models. These semantic annotations are used as a basis for designing adequate search and navigation structures. In addition, the annotations will be used as the basis for analysis and comparison of process models in the repository. The annotation model is based on well-established business frameworks, existing process classification schemes, organizational theories and other perspectives of a business process. The novelty of the annotation model can be found in the conceptualization of the business framework (REA (Geerts & McCarthy, 2000; Dunn et al, 2005)), the conceptualization of existing process classification schemes (Porter's Value Chain (Porter, 2008), the Open-EDI framework (UN/CEFACT, 2003)) and the conceptualization of enterprise modelling concepts (Huat Lim et al, 1997; Fox et al, 1996).

3. A business process relationship meta-model for identifying and defining the relationship between process models in the repository.

As the main contribution of this research, the business process relationship meta-model is developed for identifying and defining the relationship of models in the repository. The purpose of the meta-model is to enable users to identify and define the relationship between process models in the repository, which serves as the navigation mechanism. Process relationships serve as a vehicle for depicting the link between process models and thus a means for helping users to navigate the repository. These relationship definitions are used as a main contributing element for designing navigation structures between related processes. The meta-model is based on existing and well-established process relationships and process-assets and asset-processes archetypes we have developed as a method to find all processes that exist in an enterprise. The novelty of the process relationship meta-model can be found in the conceptualization of the components of an enterprise (assets, sensor and processes) (Bider et al, 2011), and the

conceptualization of the process-assets and asset-processes archetypes (Bider et al, 2012).

4. An architecture of an open and language-independent process model repository with an efficient retrieval system to process model reuse.

Another contribution of this thesis is the architecture of a process model repository. Architecture provides the foundation on which systems are built. From the system design perspective, by designing the architecture, we have structured the solution to solve the problems of existing repositories. The purpose of designing the architecture of the process model repository is to bridge the existing gap by providing the basis for developing a repository system that will increase the probability of process model reuse. The architectural design includes descriptions of different components of the system and how they address the existing challenges. The main goal is not to suggest the best possible architecture, but to show that a good enough architecture can be designed based on the known architectural principles and knowledge sources. The principles and the knowledge sources were chosen based on their fitness for the task at hand, the main requirement being that they can be integrated in a reasonable whole that can be used for developing a repository.

1.5 Publications

This thesis builds on the papers that have been accepted and published in the proceedings of various international conferences and journals. These publications are listed below. The next section describes the structure of the thesis showing how the chapters and the papers are organized and presented.

Paper I

Elias, M. & Johannesson, P. (2012b) A survey of process model reuse repositories. In: Dua, S., Gangopadhyay, A., Thulasiraman, P., Straccia, U., Shepherd, M., & Stein, B. (eds.) *Information Systems, Technology and Management*. Grenoble, France, Springer Berlin Heidelberg, pp. 64–76.

This paper presents a survey of existing process model repositories. The main goal of this paper is to identify the main challenges that limit existing process model repositories from supporting the reuse of process models.

In this paper Muri Elias is the main author and contributed to all the sections of the paper. He made the main contribution in establishing a survey protocol, identifying existing repositories through searching based on the survey protocol, and then reviewing and analysing the repositories based on the established requirements. Muri's contribution to the paper corresponds to about 80 percent.

Paper II

Shahzad, K., Elias, M. & Johannesson, P. (2010) Requirements for a business process model repository: A stakeholders' perspective. In: Abramowicz, W. & Tolksdorf, R. (eds.) *Business Information Systems*. Berlin Heidelberg, Springer, pp. 158–170.

This paper presents a set of stakeholder requirements for a process model repository. The main goal of this paper was to establish, from key stakeholders (practitioners and researchers), a set of requirements that a repository must possess to increase the probability of process model reuse.

In this paper Mhuri Elias is the second author and contributed to all the sections of the paper. He made the main contribution in conducting the exploratory study, transcribing and analysing the results of the study and defining the requirement propositions, which were the basis for the validated requirements. Mhuri's contribution to the paper corresponds to more than 40 percent.

Paper III

Elias, M. & Johannesson, P. (2013) A context-based process semantic annotation model for a process model repository. *Business Process Management Journal*, 19, 404–430.

This paper presents a context-based process semantic annotation model (CPSAM) for annotating process models in the repository. The main goal of this paper was to develop a semantic annotation model that can be used to annotate process models with information that can facilitate searching, navigation and understanding of process models.

In this paper Mhuri Elias is the first author and contributed to all the sections of the paper. He made the main contribution in reviewing existing literature to identify the potential annotation elements that formed the basis for constructing the model, conducted a confirmatory study to validate the elements and then constructed the annotation model. Mhuri's contribution to the paper corresponds to about 80 percent.

Paper IV

Elias, M., Shahzad, K. & Johannesson, P. (2010) A business process metadata model for a process model repository. In: Bider, I., Halpin, T., Krogstie, J., Nurcan, S., Proper, E., Schmidt, R., & Ukor, R. (eds.) *Enterprise, Business-Process and Information Systems Modeling*. Berlin Heidelberg, Springer, pp. 287–300.

This paper presents a business process metadata model for annotating process models in the repository. The aim of this paper was to develop and

evaluate a metadata model that can be used to annotate process models with information that can facilitate searching, navigation and understanding of process models.

In this paper Mturi Elias is the first author and contributed to all the sections of the paper. He made the main contribution in reviewing existing literature to identify the metadata elements that formed the basis for constructing the model, conducted a confirmatory study to validate the elements and then constructed the metadata model. Mturi also designed and conducted a controlled experiment to evaluate the model. Mturi's contribution to the paper corresponds to about 70 percent.

Paper V

Elias, M. & Johannesson, P. (2012a) An empirical assessment of the effect of context-based semantic annotation on process model discovery. In: Bajec, M. & Eder, J. (eds.) *Advanced Information Systems Engineering Workshops*. Berlin Heidelberg, Springer, pp. 366–382.

This paper presents an empirical assessment of the effect of the context-based process semantic annotation model (CPSAM) on process model discovery. The main goal of this paper is to test whether process annotation based on the CPSAM can improve searching, navigation and understanding of process models stored in a repository.

In this paper Mturi Elias is the main author and contributed to all the sections of the paper. He made the main contribution in designing and conducting the experiment as well as analysing the results of the experiments. Mturi's contribution to the paper corresponds to about 80 percent.

Paper VI

Bider, I., Perjons, E. & Elias, M. (2012) Untangling the dynamic structure of an enterprise by applying a fractal approach to business processes. In: Sandkuhl, K., Seigerroth, U. & Stirna, J. (eds.) *The Practice of Enterprise Modeling*. Berlin Heidelberg, Springer, pp. 61–75.

This paper presents a fractal approach to business processes as a method to untangling the dynamic structure of an enterprise. The main goal of this paper is to develop a procedure for identifying all processes that exist in an enterprise as well as their interconnections.

In this paper Mturi Elias is the third author and contributed to all the sections of the paper. He made the main contribution in designing and conducting the interviews, transcribing and analysis of the interview. Mturi's contribution to the paper corresponds to more than 30 percent.

Paper VII

Elias, M., Bider, I. & Johannesson, P. (2014) Using fractal process-asset model to design the process architecture of an enterprise: Experience report. In: Bider, I., Gaaloul, K., Krogstie, J., Nurcan, S., Proper, H., Schmidt, R., & Soffer, P. (eds.) *Enterprise, Business-Process and Information Systems Modeling*. Berlin Heidelberg, Springer, pp. 287–301.

This paper presents a report on a project of applying the process-assets and asset-processes for designing process architecture of an enterprise. The main goal of this paper was to validate the archetypes in a real-world case study from a higher education institution.

In this paper Mturi Elias is the main author and contributed to all the sections of this paper. He made the main contribution in designing and conducting the case study as well as analysing the results. Mturi's contribution to the paper corresponds to about 90 percent.

Other Publications

Below is the list of publications related to the theme of the thesis but not used in the thesis.

Paper VIII

Shahzad, K., Elias, M. & Johannesson, P. (2009) Towards Cross Language Process Model Reuse – A Language Independent Representation of Process Models. In: Persson, A. & Stirna, J. (eds) *The Practice of Enterprise Modeling*. Springer Berlin Heidelberg, pp. 176–190.

Paper IX

Elias, M. & Shahzad, K. (2010) Using multi-criteria decision making to choose process representation format for a process repository. In: Abramowicz, W., Tolksdorf, R. & Węcel, K. (eds.) *Business Information Systems Workshops*. Berlin Heidelberg, Springer, pp. 19–24.

Paper X

Elias, M. (2010) Towards a Universal Business Process Model Repository for Process Model Reuse. Department of Computer and Systems Sciences, Stockholm University.

1.6 Structure of the Thesis

This thesis is organized into ten chapters and its structure is shown in Figure 1. The chapters build on the papers listed in the previous section. In this introduction chapter, we have explained the motivation of the work, described the problems, listed the research goals, and presented our major

contributions and publications. The rest of the thesis is organized as follows. Chapter 2 provides a background, context and relevant work survey for this research. Chapter 3 provides the research design. Chapter 4 presents the requirements for the process model repository. Chapter 5 presents a semantic annotation model for annotating process models, while Chapter 6 describes a prototype implementation of the repository that implements a semantic annotation model. Chapter 7 evaluates the annotation model. Chapter 8 presents a meta-model for identifying and defining the relationship between process models in the repository. Chapter 9 presents the architecture of the repository that fulfils the defined requirements. Finally, conclusions and further work are discussed in Chapter 10. The appendices are not shown in the structure shown in Figure 1, and nor is Paper I, which is related to Appendix A.1.

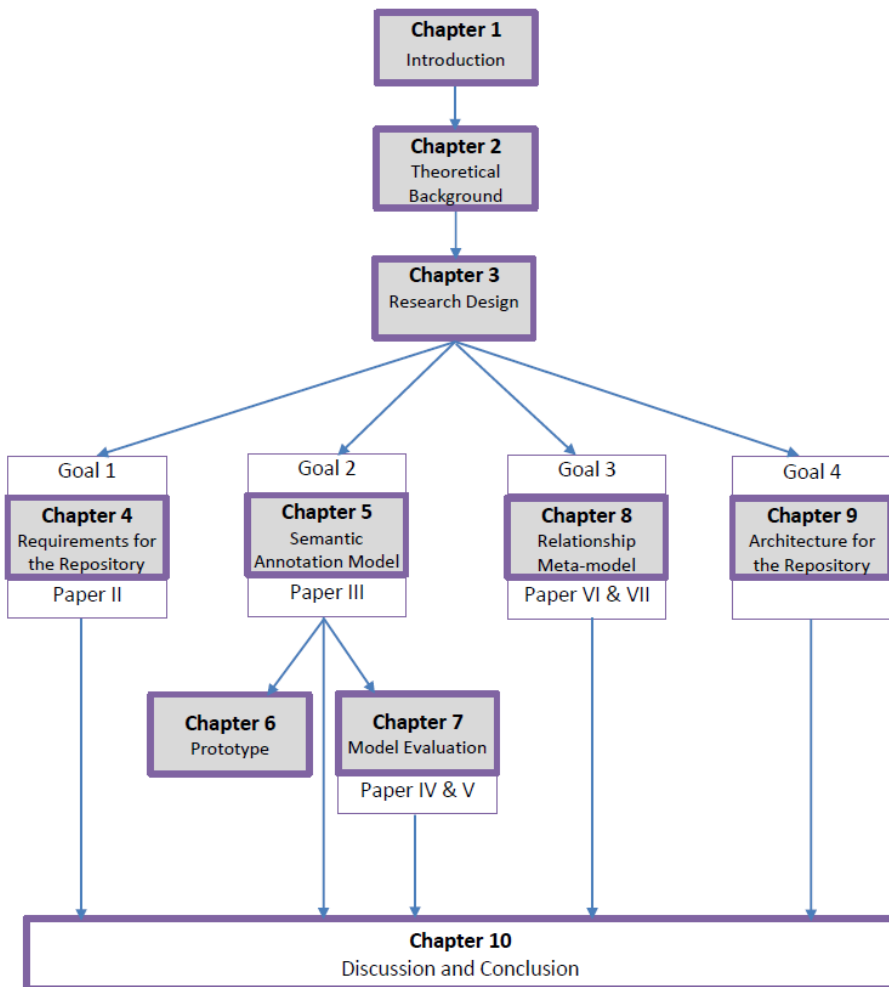


Figure 1: Structure of the thesis.

Chapter 2 – *Theoretical and Technical Foundations* establishes basic concepts, theoretical and technical background and the context relevant to all areas that influence our research. The chapter has five subsections. First, an overview of the chapter is given. Second, an overview of business process management (BPM) and business process modelling as the core component of BPM is discussed. Third, the definition and design considerations of the business process repository are provided. Fourth, the software and system architecture concepts are described. Fifth, the conceptual framework of the thesis is presented. Sixth, the summary and discussions of this chapter are presented.

Chapter 3 – *Research Design* introduces the methodological foundation and the research process of this thesis. The chapter has two subsections. First, an overview of the research paradigm we have followed in this thesis is given. Second, the research process for each research goal is described.

Chapter 4 – *Requirements for a Process Model Repository* introduces the requirements that a repository must fulfil in order to increase the probability of process model reuse. Therefore, it mainly addresses research goal 1. The chapter has six subsections. First, an overview of the chapter is given. Second, the requirements elicitation from stakeholders is described. Third, the requirements elicitation from the literature is described. Fourth, the requirements specifications and justifications are given. Fifth, some of the related works are discussed. Sixth, the summary and discussions of this chapter are presented.

Chapter 5 – *A Context-based Process Semantic Annotation Model* presents the semantic annotation model for semantically annotating process models in the repository to facilitate searching of process models, navigating the repository and understanding of process models. Therefore, it mainly addresses research goal 2. The chapter has seven subsections. First, an overview of the chapter is given. Second, the requirements for the annotation model are defined. Third, the annotation development process is described. Fourth, a context-based process semantic annotation model is described. Fifth, a demonstration of how the annotation model is applied is presented. Sixth, some of the related works are discussed. Seventh, the summary and discussions of this chapter are presented.

Chapter 6 – *Prototype implementation* presents a prototype of the Semantic Annotation Tool. The chapter has three subsections. First, an overview of the chapter is given. Second, the prototype system is described. Third, the summary and discussions of this chapter are presented.

Chapter 7 – *Evaluation of the Semantic Annotation Model* presents the evaluation we have performed to test the correctness, consistency, performance and user's perception of the annotation model. The chapter has

five subsections. First, an overview of the chapter is given. Second, the evaluation framework is described. Third, the evaluation of correctness and consistency is presented. Fourth, the evaluation of user performance and perception is presented. Fifth, the summary and discussions of this chapter are presented.

Chapter 8 – *Business Process Relationship: The Meta-model* presents a model for identifying and defining the relationship between business processes in the repository to improve the usage of process models. Therefore, it mainly addresses research goal 3. The chapter has eight subsections. First, an overview of the chapter is given. Second, the requirements for the relationship meta-model are defined. Third, the meta-model development process is described. Fourth, the process-assets and asset-processes archetypes, which form the basis for the meta-model, are described. Fifth, the validation of the archetypes through a case study is presented. Sixth, the meta-model is presented. Seventh, the evaluation of the meta-model is discussed. Eighth, the summary and discussions of this chapter are presented.

Chapter 9 – *Architecture for the Process Model Repository* presents the architecture of the repository design to meet the requirements presented in Chapter 4. Therefore, it mainly addresses research goal 4. The chapter has seven subsections. First, an overview of the chapter is given. Second, the architecture design and development process are described. Third, the architecture specification is given. Fifth, the evaluation of the architecture is presented. Sixth, some of the related works are discussed. Seventh, the summary and discussions of this chapter are presented.

Chapter 10 – *Conclusion and Future Work* summarizes the work, and outlines future work to point out the possible improvements and the interesting directions of further research on process model repositories in business process management. The chapter has four subsections. First, a review of the research goals and findings of the thesis are given. Second, a summary of the contributions is presented. Third, the limitations of this research are discussed. Fourth, future research directions are presented.

1.7 Conceptual Framework Used in the Thesis

A conceptual framework used in this thesis is presented in this section. The framework, shown in Figure 2, provides the conceptual foundation for the work presented in this thesis. Below we define the key concepts and their relationships.

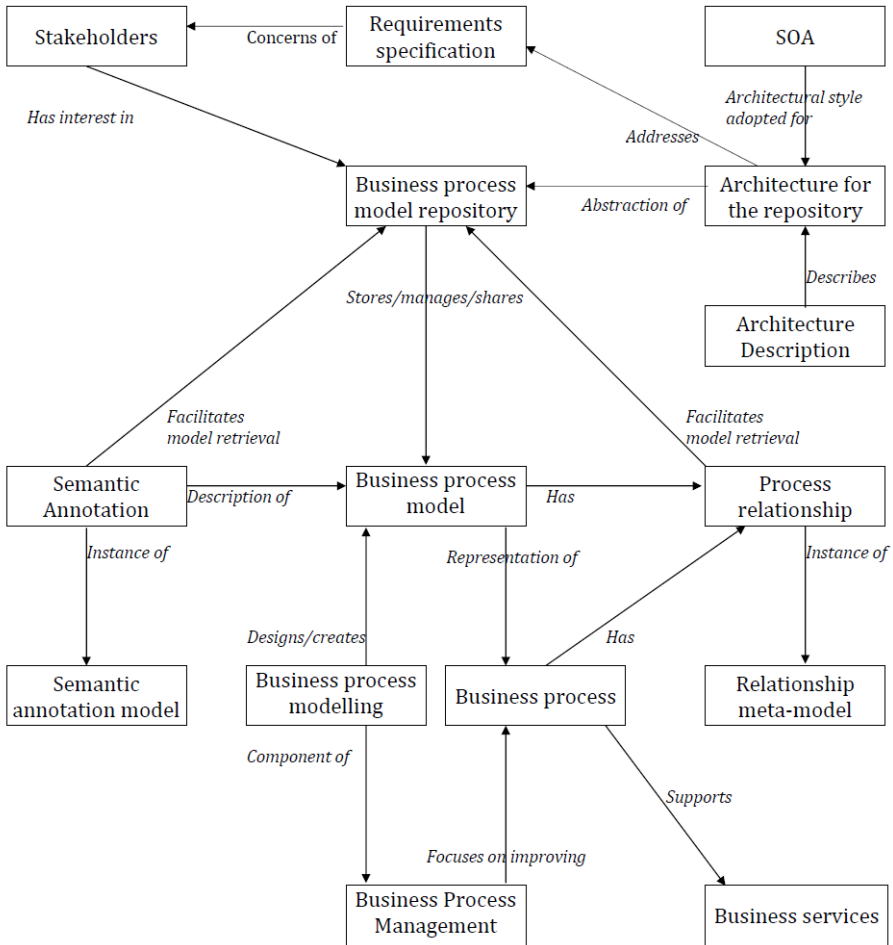


Figure 2: A conceptual framework of the key concepts used and their relationships.

Annotation: Annotation is metadata attached to process models. It is used to enrich a process model’s information and can be in the form of text descriptions. (Wikipedia Contributors, 2014a; Liao et al, 2011).

Annotation element: Annotation element is a unit of data that describes a specific property or characteristic of a business process.

Architectural description (AD): AD is a collection of products to document architecture (ISO/IEC/(IEEE), 2011).

Architecture: Architecture is the basic organization of a system embodied in its components and the relationships between them as well as the environment, and the principles guiding its design and evolution (IEEE, 2007).

Architecture view: A view is a depiction of one or more structural aspects of an architecture that show how the architecture addresses one or more concerns of its stakeholders (IEEE, 2007; ISO/IEC/(IEEE), 2011).

Architecture viewpoint: A viewpoint is a collection of patterns, templates and conventions for constructing one type of view (IEEE, 2007; ISO/IEC/(IEEE), 2011).

Business process: A business process is a chain of activities, structured or unstructured, performed to produce a specific service or product for a particular customer or customers (Davenport, 1993; Hammer & Champy, 1993; Wikipedia Contributors, 2014b).

Business process management (BPM): BPM is a management approach that focuses on improving organization performance by managing and optimizing its business processes (Panagacos, 2012; Dumas et al, 2013; Weske, 2012).

Business process model: A process model is an abstract description of process that represents process elements that can be enacted by a human or machine (Curtis et al, 1992).

Business process modelling: Process modelling is an activity in which the processes of an enterprise are represented using process models. It involves depicting, capturing and understanding the organization's operations and its information architecture (Danesh & Kock, 2005).

Process annotation model: A process annotation model is a conceptual model that constitutes a set of elements for describing or annotating a business process or process model (Lin, 2008).

Repository: A repository is a specialized, extensible database application that adds value to a database system by being customized to a particular domain (Bernstein & Dayal, 1994; Embley & Goldstein, 1997).

Requirements: A requirement, in systems and software engineering, is a statement that identifies a necessary capability, attribute, characteristic or quality of a system for it to have value and utility to its stakeholders. A **requirement specification** is an explicit set of requirements to be satisfied by a design, product or service.

Semantic annotation: A semantic annotation is a semantic description assigned to an entity, such as a process model (Liao et al, 2011).

Service-Oriented Architecture (SOA): SOA is an architectural style for designing information systems where the goal is to achieve loose coupling among interacting system components (Erl, 2005).

Stakeholder: A stakeholder is an individual with an interest or concern in a system.

2 Theoretical and Technical Foundations

2.1 Overview

This chapter positions the research in this thesis by establishing basic concepts, theoretical and technical background and the context relevant to all areas that influence our research. Since the problem we are addressing is multifaceted, our literature review spans different fields. The discussion ranges from business process management, business process modelling and enterprise modelling to software and systems engineering. The chapter begins by providing an overview of business process management (BPM), where a detailed discussion of what a business process is, with examples, is provided. The overview includes discussion of the foundation of BPM and outlines the BPM life cycle. The chapter provides an overview of business process modelling and the concepts and relevant issues about business process modelling. We explain how process modelling influences most activities of the BPM life cycle. We also discuss different perspectives for comprehensively representing a process model, and various process modelling approaches. The challenges of modelling business process from scratch are highlighted along with the need for reusing process models. The chapter discusses the research target – a business process repository – as an approach to supporting process model reuse by enabling storing and sharing of process models. The repository design considerations are discussed and related to our research goals. The chapter concludes with some key remarks in relation to the research targets.

2.2 Business Process Management

2.2.1 Business Processes

Every organization performs a chain of events, activities and decisions that ultimately add value to the organization and its customers. These chains of events, activities and decisions are called “business processes” (Dumas et al, 2013).

As defined by Davenport (1993), a business process is “a structured, measured set of activities designed to produce a specific output for a particular customer or market”. The emphasis is on “how” work is done as opposed to a product focus’s emphasis on “what”. Consequently, a process is

a specific ordering of activities across time, with a start and an end, and clearly defined inputs and outputs (Davenport, 1993). Fundamentally, processes are the structures used by an organization to perform what is needed to produce value for its customers.

The same “value” in relation to process outcome is promoted by Hammer and Champy (1993) as they define a business process as “a collection of activities that takes one or more kinds of input and creates an output that is of value to the customer”. This definition can be seen as a subset of Davenport. Another key characteristic of a business process is that of the transformation of input into output. Johansson et al (1993) define a business process as “a set of linked activities that take an input and transform it to create an output”. Ideally, the transformation that occurs should add value to the input and create a valuable and effective output to the recipient.

Technically, a business process is a structured set of activities that takes an input and transforms it into a more valuable and effective service or product (serve a particular goal), as an output, for a particular customer or

Table 1: Business process definitions drawn from the literature

<i>Definitions</i>	<i>Characteristics</i>
A <i>business process</i> is a structured, measured set of activities designed to produce a specific output for a particular customer or market. It implies a strong emphasis on how work is done within an organization, in contrast to a product focus’s emphasis on what. A process is thus a specific ordering of work activities across time and space, with a beginning and an end, and clearly defined inputs and outputs: a structure for action (Davenport, 1993).	<ul style="list-style-type: none"> • Activities ordered across time and space • Input and output • Has a customer • Organization
A <i>business process</i> is a collection of activities that takes one or more kinds of input and creates an output that is of value to the customer (Hammer & Champy, 1993).	<ul style="list-style-type: none"> • Collection of activities • Input and output • Creates value • Has a customer
A <i>business process</i> is a set of linked activities that take an input and transform it to create an output (Johansson et al, 1993).	<ul style="list-style-type: none"> • A set of activities • Transformation • Input and output
A <i>business process</i> is “a series of steps designed to produce a product or service. Most processes are cross-functional, spanning the ‘white space’ between the boxes on the organization chart. Some processes result in a product or service that is received by an organization's external customer. We call these primary processes. Other processes produce products that are invisible to the external customer but essential to the effective management of the business. We call these support processes” (Rummler & Brache, 1995).	<ul style="list-style-type: none"> • A series of steps • Produces a product or service • Crosses functional boundaries • Primary or support process

customers. In fact, a business process is not only a structured set of activities; it is also characterized by having unstructured set of activities (Swenson et al, 2012; Herrmann & Kurz, 2011). The activities of a business process may be performed by people manually or with the support of information systems and are completed either sequentially or in parallel (Weske, 2007). Also, there are activities of a business process that can be automatically enacted by information systems, without any human participation (Weske, 2007). Table 1 summarizes several definitions drawn from the literature and extracts the main characteristics of a process.

Thus, from the above definitions we can resolve that a business process has the following characteristics:

- A business process is comprised of a set of activities.
- Activities take an input and transform it to a valuable output, which is either a product or a service.
- The transformation is aimed at creating value for customers.
- A business process has its customers. The customers may be internal or external to the organization.
- Activities are performed by actors, which may be human or machines.
- A business process realizes some business goals.
- A business process often involves more than one organizational unit, which are responsible for a whole process.

Examples of business processes (Dumas et al, 2013), are:

- *Quote-to-order*. This type of process typically precedes an order-to-cash process. It starts from the point when a supplier receives a “Request for Quote” (RFQ) from a customer and ends when the customer in question places a purchase order based on the received quote. The order-to-cash process takes the relay from that point on. The combination of a quote-to-order and the corresponding order-to-cash process is called a “quote-to-cash process”.
- *Application-to-approval*. This type of process starts when someone applies for a benefit or privilege and ends when the benefit or privilege in question is either granted or denied. This type of process is common in government agencies, for example when a citizen applies for a building permit or when a businessman applies for a permit to open a business (e.g. a restaurant). Another process that falls into this category is the admissions process in a university, which starts when a student applies for admission into a degree. Yet another example is the process for approval of holiday or special leave requests in a company.

An organization’s success hinges to a large extent on how well its business processes are designed and performed (Weske, 2007). Business process management (BPM) is a management approach that focuses on improving organization performance by managing and optimizing its business processes (Panagacos, 2012; Dumas et al, 2013). It is a

holistic management approach that fosters business effectiveness and efficiency as it strives for innovation, flexibility and integration with technology (Chavan & Lal, 2012). Compared to the functionally focused management approach, the BPM approach enables organizations to be more effective and efficient, and more capable of change (Ko, 2009; Weske, 2012).

The BPM discipline combines knowledge from business administration and information technology and applies this to operational business processes (Weske, 2007; Weske, 2012). Members of business administration are interested in improving the company's operational performance. The important aspects of BPM from a business administration point of view are increasing customer satisfaction, reducing the cost of doing business, and establishing new products and services at low cost (Weske, 2007; Weske, 2012). On the other hand, the computer science community is interested in investigating structural properties of business processes and providing robust and scalable software systems (Weske, 2007; Weske, 2012).

2.2.2 BPM Foundation

Business process management (BPM) is rooted in both management science and computer science (van der Aalst, 2013). As a result it is difficult to pinpoint its starting point.

A good starting point for establishing the road to BPM is from a management concept referred to as business process re-engineering (BPR) (Hammer & Champy, 1993). In this concept a new way of organizing companies on the basis of business processes was proposed. Following their campaign for the radical redesign of business processes, many companies initiated BPR projects to review and redesign their processes. However, in the late 1990s, the interest in BPR deteriorated and many enterprises terminated their BPR projects and stopped supporting further BPR initiatives (Jeston & Nelis, 2014).

The emergence of BPM can be seen from two key renewed ideas behind BPR (Dumas et al, 2013). The first is the revelations from the empirical studies, which showed that process-oriented organizations did better than non-process-oriented organizations (Dumas et al, 2013). As a result of the confirmation of this picture by the follow-up studies, the credibility of the process-oriented concept was renewed.

The second is the technological development. Following BPR in the late 1990s, enterprise resource planning (ERP) systems and workflow management (WFM) systems gained organizational focus (Dumas et al, 2013). ERP is a business management software, usually a suite of integrated applications, used for managing the business and automating the back office functions related to services, technology and human resources (Sumner, 2004; Leon, 2012). It integrates all components of an operation, ranging from product planning, development and manufacturing to marketing and sales. On the other hand, WFM systems are systems that distribute work to

various actors in a company on the basis of process models (Dumas et al, 2013).

BPM can be perceived as an extension of workflow management (WFM) (van der Aalst, 2013). While the primary focus of WFM is to automate business processes (van Der Aalst & Van Hee, 2004), the scope of BPM is broader – ranging from process automation and analysis to organization of work and operations management. On the one hand, BPM is aimed at improving business processes, perhaps without the use of technologies (Weske, 2012). For example, management may get ideas on how to reduce costs by modelling and analysing a business process using simulation.

2.2.2.1 BPM Systems

BPM is often related to business process management systems (BPMS) – a software that supports modelling, analysis and enactment of business processes (van der Aalst et al, 2003b; Chang, 2005). In the BPM community, it is widely agreed that the core of a BPM system is the functionality that has been attributed traditionally to a WFM system (van der Aalst et al, 2003b).

From a workflow perspective, a BPM system is understood as mainly responsible for the automatic work allocation to authorized and qualified resources according to a predefined plan of the process (Jablonski & Bussler, 1996; Lawrence, 1997; van Der Aalst & Van Hee, 2004). This means that before a WFM system can become effective, business processes within an organization should be identified, analysed and mapped (Reijers, 2006). A BPM system extends the capabilities of the WFM systems by providing extra sophisticated build and run time diagnostic capabilities. The addition of these capabilities has led to a significantly increased potential for BPM systems to support change in organizational processes (Shaw et al., 2007).

2.2.2.2 Process-Aware Information Systems

BPM methods are not only limited to WFM and BPM systems, but extend to any Process-Aware Information System (PAIS) (van der Aalst, 2013). PAIS is a software system for managing and executing organizational processes that involve people, applications and/or information sources on the basis of process models (Dumas et al, 2005). According to this definition, a PAIS is an information system for supporting business processes. It is mainly configured by a business process model (Van der Aalst & Stahl, 2011). PAISs include conventional WFM systems and contemporary BPM systems. They also include systems for supporting specific processes or providing more flexibility (Dumas et al, 2005; Ghani et al, 2008). Examples of such systems are case-handling systems, ERP systems, customer relationship management (CRM) systems, etc.

2.2.2.3 Adaptive Case Management

BPM systems are well-established tools that increase business productivity. However, they have only been well suited to supporting structured business processes where everything about the process can be defined in advance

(Herrmann & Kurz, 2011; Swenson et al, 2012). Unlike structured processes, in *unstructured* business processes, such as knowledge-intensive work, the order of activities cannot be predetermined in advance (Swenson et al, 2012). An example of an unstructured business process is an insurance claim, which relies on unstructured interactions among experts who assess the claim in the field before moving to a more structured series of activities for processing the claim within the insurance company.

Change and unpredictability management is an important part of doing business. Organizations that have managed to control this flexibility in their business processes see more value and become more competitive (Khanna, 2013). Adaptive Case Management (ACM) is a new BPM concept, aimed at flexibly supporting these knowledge-intensive, less structured or unstructured business processes, along with a corresponding software system (Herrmann & Kurz, 2011).

A case is commonly an extended and collaborative process. The case processing entails coordination of knowledge, content, correspondence and human tasks (Khanna, 2013). In most of the processes, the case requires adherence to business rules and policies as well as regulatory requirements. It is called adaptive because the process execution path cannot be fully predefined and may have to adapt to the specific requirements of every case (Khanna, 2013). At various stages in the process, human judgment is required to determine the next step in the processing of the case. During the time of execution, external events and case types may alter the processing.

Adaptive case management provides them with a lot of flexibility while ensuring the visibility and audit tracking of each case.

In this section we have provided the BPM foundation and showed how it is related to WFM, ACM and PAIS in general. In the following section, we provide an overview of the BPM life cycle that builds a stage for business process modelling and business process models.

2.2.3 The BPM Life Cycle

There are many views of the BPM life cycle (van der Aalst et al, 2003b; Havey, 2005; Netjes et al, 2006; Weske, 2007; Dumas et al, 2013). Some of the BPM life cycle models such as the one proposed by van der Aalst et al (2003b) are strongly associated with process automation using BPM technologies, whereas in reality BPM can be implemented without adopting any BPM technologies (Jeston & Nelis, 2014). In this thesis we have adopted the one proposed by Dumas et al (2013) (see Figure 3) because it covers both aspects of BPM. According to them, activities of business process management (BPM) can be grouped into six phases: identification, discovery, analysis, redesign, implementation, and monitoring and control (Dumas et al, 2013).

The first phase is *process identification*. This phase includes activities related to the identification of processes relevant to the business problem being addressed. It also includes the delimitation of the scope of these

processes as well as identification of the relationship between them. The outcome of this phase is a process architecture that provides a holistic view of the processes in an organization and how they are related.

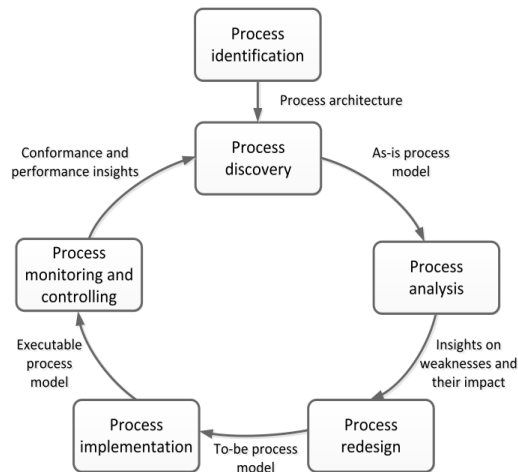


Figure 3: BPM life cycle, taken from (Dumas et al, 2013).

The second phase is *process discovery*. Once the processes and the performance measures are identified, the next step is to understand these business processes in detail. This phase is called “process discovery”; it is also referred to as “as-is process modelling”. In this phase, the current state of each of the relevant processes is documented, usually in the form of one or several as-is process models. The as-is process models are the reflection of how the work is done in the organization.

The third phase is *process analysis*. This phase includes activities related to identification and documentation of problems associated with the as-is process. Accordingly, if possible the identified problems are quantified using performance measures. The output of this phase is a structured collection of problems.

The fourth phase is *process redesign*. This phase includes activities for identification of changes to the process that would help to address the problems identified during the analysis phase. Several change options are examined and compared based on the chosen performance targets. Ultimately, the best change options are combined, leading to a redesigned process. The output of this phase is usually a set of to-be business process models.

The fifth phase is *process implementation*. In this phase, the changes needed to migrate from the as-is process to the to-be process are implemented. The implementation ranges from organizational change management to process automation. Organizational change management includes activities required to change the way of working, whereas process

automation refers to the development and deployment of IT systems that support the to-be process.

The last phase is *process monitoring and controlling*. In this phase, relevant data related to the implemented process are collected and analysed to establish how well the process is performing. Bottlenecks, recurrent errors or deviations with respect to the intended behaviour are identified and corrective actions are undertaken. New issues may then arise, in the same or in other processes, requiring the cycle to be repeated on a continuous basis.

With these BPM life cycle phases we are able to see how business process models and business process modelling are fundamental components in the implementation of BPM.

2.3 Business Process Modelling

A model is an abstract depiction of reality that eliminates much of the infinite detail of the world (Curtis et al, 1992). The primary drive of a model is to reduce the complexity of comprehending a phenomenon by removing unnecessary details (Curtis et al, 1992). Therefore, a model uncovers what its designer believes is essential in comprehending the phenomenon modelled.

Curtis et al (1992) define a business process model as “an abstract description of an actual or proposed process that represents selected process elements that are considered important to the purpose of the model and can be enacted by a human or machine”. A process element is any component of a process such as an activity and data. A business process model is a simplified, abstract representation of a business process.

The activity of representing processes of an enterprise using process models is called “business process modelling”. It involves depicting, capturing and understanding the organization’s operations and its information architecture (Danesh & Kock, 2005). Process modelling is widely considered to be a critical component in successful business process management (BPM) (Dumas et al, 2013; Weske, 2007). Notably, process modelling is essential within a BPM life cycle. It plays important roles in the process discovery, analysis, design and implementation phases within BPM (Dumas et al, 2013).

- *Process discovery*. During the process discovery phase, process modelling is applied to map out current (or “as-is”) processes, of an organization, to create a baseline for process improvements.
- *Process analysis*. Process models produced during the process discovery phase are used in the analysis phase as a torch in order to shed more light on the process, expose the main causes of problems in the process and discover possible problem solutions.
- *Process design*. In the design phase, process modelling is applied to create future (or “to-be”) processes, which would address the issues identified in the as-is process models. Process models are used in the

design phase to help define the core processes of a business covering activities and their dependencies, data flow, roles and the actors involved, and established goals.

- *Process implementation.* Process models are used in the implementation phase, in which new business processes are put into actual use. The implementation can be done through a special kind of process automation, which uses IT systems such as business process management systems (BPMSs) to help coordinate, control and communicate process execution. BPMSs exploit *process models* to coordinate business processes.

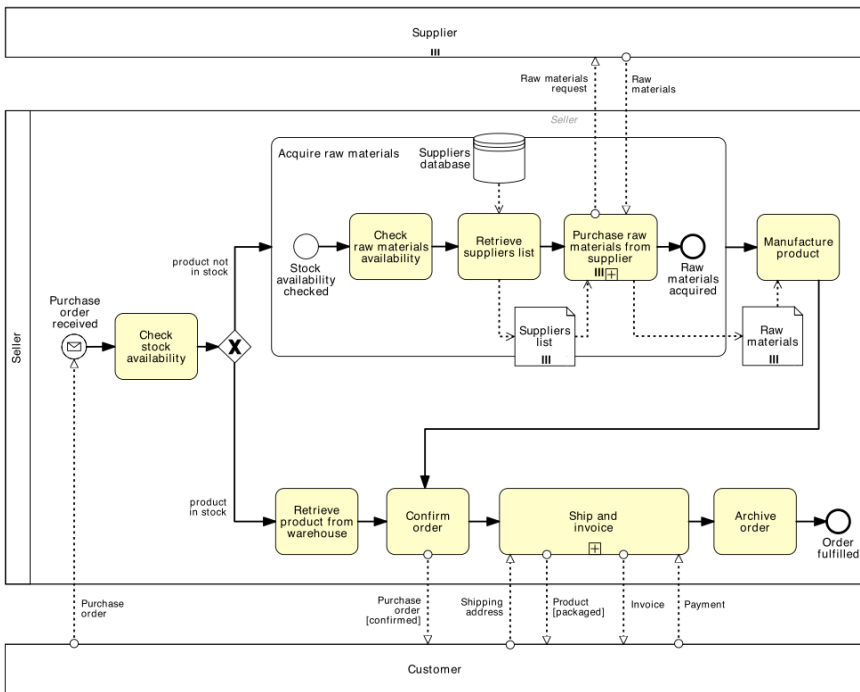


Figure 4: An order fulfilment process model example.

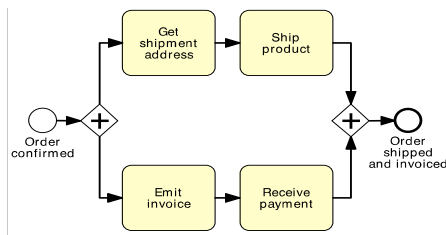


Figure 5: Ship and invoice subprocess of the order fulfilment process

Besides the above three functions, process modelling can possess the analysis capability to facilitate process monitoring and evaluation.

Process modelling often uses modelling language for representing processes with flow chart-like diagrams that can be easily comprehended by both IT and business personnel. Such a modelling language is often called “graphical modelling language”. Some of the widely used graphical modelling languages include Event-driven Process Chain (EPC) (Scheer, 2000), Business Process Model and Notation (BPMN) (Object Management Group, 2006), Unified Modelling Language (UML) (Object Management Group, 2004) and Petri net (Petri, 1962). A process modelling language can also be a set of standardized textual keywords or markup language accompanied by parameters, for instance Business Process Modelling Language (BPML) (Arkin, 2002), Business Process Execution Language (BPEL) (Andrews et al, 2003) and EPML (Mendling & Nüttgens, 2006). Graphical modelling languages usually facilitate readability of models whilst textual modelling languages enable models to be machine-interpretable and tool-interchangeable (Ottensooer et al, 2012).

Figure 4 illustrates an example of an order fulfilment process model using BPMN.

2.3.1 Process Representation Perspectives

Describing a business process requires integration of many forms of information. Some forms of information that people typically want to obtain from a process model are what is going to be done, who is going to do it, when and where will it be done, how and why will it be done, and who is dependent on its being done (Curtis et al, 1992). Process modelling languages vary in the level to which their constructs specify the information that answers these different questions (Curtis et al, 1992). Usually modelling languages present one or more different perspectives related to these questions. The most commonly represented perspectives are *functional*, *behavioural*, *organizational* and *informational* as proposed by Curtis et al (1992). Korherr (2008) suggests an additional perspective called the *business process context perspective*.

2.3.1.1 Functional Perspective

The *functional* perspective focuses on what activities are performed in a business process and how these transform resources from input to output. The activities of a process can either be atomic or composite activities, which are recursively refined by activities. The functional perspective of a business process addresses the question “What is to be done?”

As an example, consider the order fulfilment process model shown in Figure 4. The functional perspective of this process includes description of atomic activities such as ‘check stock availability’ and ‘check material availability’. The model also consists of the composite activities, also referred to as sub processes, purchase raw materials from supplier, and ship and invoice shown in Figure 5.

2.3.1.2 Behavioural Perspective

The *behavioural* perspective of a business process focuses on the timing of activities – when activities are to be performed – as well as the ordering and flow of activities and data. A flow can be either a *control flow* or a *data flow*. A *control flow* expresses when an activity is to be performed in relation to others, whereas a *data flow* defines the flow of information resource among activities and message exchanges among actors. Process specifications rely on control flow patterns for specification of coordination rules among activities. The control flow patterns include sequence flow, parallel execution (AND-Split), conditional branching (OR-Split, XOR-Split) and synchronization (AND-Join, OR-Join, XOR-Join, N-Out-of-M-Join). These patterns make it possible to specify decisions made according to certain business rules. The behavioural perspective addresses the question “When and how will it be done?”

As an example, the order fulfilment process shown in Figure 4 includes the sequence flow such as the arrow that connects and controls the dependencies between “Retrieve product from warehouse” and “Confirm order” so that the latter cannot start before the former is finished. It also includes the conditional branching such as the XOR-Split that decides whether a product is in the stock. If the product is in stock, the seller proceeds to retrieve the product from the warehouse. If the product is not in stock, the seller proceeds to acquire raw materials for product manufacturing. Also, an example of the parallel split included is the AND-Split when the order is confirmed in Figure 5. After concurrent execution the two branches are joined by synchronization (AND-Join).

2.3.1.3 Informational Perspective

The *informational* perspective defines informational entities produced, used, modified or exchanged during the execution of a business process (Curtis et al, 1992). The informational entities can either be data, artefacts, products or objects involved in a particular activity. Information entities make it possible to describe data or messages that are the output of one activity and show how they affect the execution of the following activity.

In the order fulfilment process shown in Figure 4, Supplier list, as a data object, is an information entity. It is an output of the “Retrieve suppliers list” activity and an input to a subprocess, “Purchase raw material from supplier”. Also, purchase order and invoice are data objects whereas supplier database is an artefact.

2.3.1.4 Organizational Perspective

The *organizational* perspective defines the allocation of the responsibility for performing activities of a business process. Mainly the focus is on the notion of the actor – organizational unit, role or software – which executes the activities of a business process. When the actor is an organizational unit, its members execute activities. A role characterizes the responsibilities and behaviour of actors. Software services as an actor can be used to automatically execute activities. The use of the organizational perspective enables dedication and control of responsibilities of the parties involved in a business process. The organizational perspective addresses the question “Where will it be done and who is going to do it?”

In the process model shown in Figure 4, the organization’s seller is the main actor that performs the order fulfilment process. The seller receives a purchase order from the customer (external actor) and checks it against the stock. If the product is not in stock, the seller proceeds to check whether the raw materials are available for the product to be manufactured. If there are no raw materials, the seller issues the raw material request from the supplier (another external actor).

2.3.1.5 Business Process Context Perspective

In addition to the four perspectives, List and Korherr (2006) introduced another perspective, namely the *business process context perspective*. The above perspectives cover the detailed sequence of a process; they do not cover important process characteristics. Therefore this perspective defines main characteristics of a business processes, such as business goals and their measures.

2.3.2 Modelling and the Need for Reuse

According to Weske (2007), business process modelling is an integration of multiple modelling subdomains – functional, data, organizational and IT landscape modelling. The functional modelling investigates and represents the units of work, in the form of activities or tasks, which are being executed in a business process context. The data modelling investigates and represents the data or information that is consumed or produced by different activities of a business process. The organization modelling investigates and represents individuals or organization units that perform different activities of a business process. The IT landscape modelling investigates and represents information systems that perform specific activities of a business process. Additional subdomains can be added depending on their relevance.

The functional, data and organizational subdomains are related to functional, informational and organizational process representation perspectives. Thus process modelling defines the glue that connects the process perspectives together. It relates activities (functional perspective) of a business process with execution constraints in order to specify the ordering and conditional execution of activities (behavioural perspective) (Weske,

2007). It also includes the data or information aspects (informational perspective) because some of the activities may depend on the data involved in a particular business process (Weske, 2007). The actors (organizational perspective) who perform the activities of the business process are also covered. For example, the order fulfilment process model shown in Figure 4 ties together the activities (i.e. check available stock, check raw material availability, etc.), and the constraints and order in which these are executed are also specified, as well as the information consumed (i.e. supplier database) and produced (i.e. supplier list) by certain activities, and the actors (i.e. seller, suppliers, customers) involved in a process execution.

Modelling includes decomposition of a business process into subprocesses and adding required process elements to the model. For example, Figure 5 depicts a decomposition of the ship and invoice subprocess of the order fulfilment process model.

Modelling of business processes requires intensive knowledge related to both (Vulcu et al., 2011) (i) the business domain that is to be modelled and (ii) the modelling language. In fact, the analysts have to define exactly what activities should be captured, their execution logic and how to specify them using a particular modelling language. Usually, the scope of modelling goes past the boundary of one process by capturing all the related processes and the business environment (Barjis, 2011). For modern enterprises, the modelling activity includes both intra-organizational and inter-organizational business processes (Weske, 2007; Barjis, 2011). Consequently, designing process models from scratch is often a complex task. Such complexity means that modelling business processes without sufficient expertise in a broad set of disciplines can result in extended modelling cycles, higher modelling costs and quality problems (Ma & Leymann, 2008; Markovic & Pereira, 2008a; Aldin & de Cesare, 2011).

While business process models are initially created for specific needs based on certain business requirements, there are some similarities between business requirements across organizations and industries (Wu et al, 2009). Furthermore, many processes share activities, subprocesses and organizational entities. Thus, the modelling experts may encounter similar modelling scenarios, within and across industries, many times, and such experience would be of great value to the organizations if documented (Havey, 2005; Aldin & de Cesare, 2011). This implies that new process models could be created by modifying existing models rather than recreating them from scratch (Lin, 2008; Aldin & de Cesare, 2011). Such reuse not only simplifies the modelling but also increases the quality and the maturity of the developed process models and reduces the cost of modelling (Awad et al., 2011).

Capturing and storing process models is the initial step towards reuse of process models. Process model repositories provide a central location to support representation, capturing, sharing and reuse of process models (Markovic & Pereira, 2008b; Hornung et al, 2009).

2.4 Business Process Model Repository

According to Bernstein and Dayal (1994), a repository is a “shared database of information about engineered artefacts produced and used by an enterprise”. However, a repository is not necessarily internal to an enterprise, but may also be open to public use. While the word “repository” has been stereotyped to mean everything from a file that stores information, to disk storage to a database of specific information, within the context of this thesis, the term “repository” is not synonymous with a database system. A repository is a specialized, extensible database application that adds value to a database system by being tailored to a specific domain such as application development (Embley & Goldstein, 1997).

A business process model repository provides a central location for collecting and sharing process knowledge (business rules, relationships, process elements, etc.) for different purposes: enabling stakeholders to retrieve process models for understanding business operations; updating, simulating and analysing process models; and reusing process models (Ma et al, 2007; Ma & Leymann, 2008; La Rosa et al, 2011a). Thus there are several instantiations of process model repositories that may in fact have different needs and different architectures.

This thesis proposes an open and language-independent process model repository – a repository that is publicly open to any potential users for capturing, sharing and reuse of process models. Central to the functions of this repository is supporting reuse of process models among different stakeholders, across organizations and industries. As defined previously, process model reuse is a systematic practice of designing new process models from a stock of existing process models (Elias & Johannesson, 2012b).

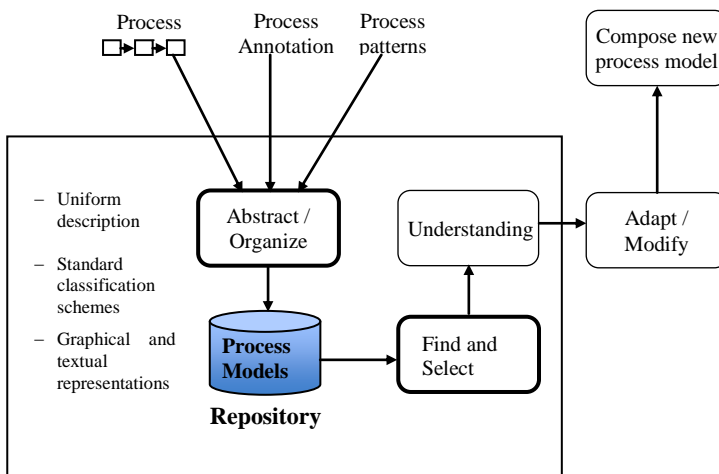


Figure 6: Process model reuse (adapted from software reuse (Constantopoulos et al., 1995))

Figure 6 offers a simplified view of the model reuse process and important functions of the repository system. For process reuse to be attractive, the overall effort to reuse must be less than the effort to create a new process model. Process model reuse involves three steps: (1) finding and selecting the process model, (2) understanding the model, and (3) adapting the model.

From the above steps of the reuse process a number of considerations should be kept in mind while designing the process model repository and its functionality. They include: (1) process knowledge and adaptability, (2) process model representation and understanding, and (3) repository structuring and retrieval.

2.4.1 Process Knowledge and Adaptability

Model adaption depends on the differences between requirements and the features offered by existing models and the skills of the user. Complete process models are the primary reusable process artefacts that should be stored and shared for designing new process models. However, several reusable process artefacts have been proposed in the field of BPM to support redesign of new process models. The decision of what process artefacts should be stored for reuse may influence the design of the repository. This is because different concepts may require different approaches to capture, represent, store and retrieve them. Some of the widely used reusable process artefacts include business process patterns (Aldin & de Cesare, 2011; Barros, 2007; Gschwind et al, 2008), process fragments (Markovic & Pereira, 2008a; Schumm et al, 2011a; Schumm et al, 2011b) and business process ontologies (Liao & Leung, 2007; Aldin & de Cesare, 2011; Markovic & Pereira, 2008a).

2.4.1.1 Business Process Patterns

According to Alexander et al (1977), “a pattern describes a problem which occurs over and over again in our environment, and the core of the solution to that problem, in such a way that you can use this solution several times without ever doing it the same way twice”. While a pattern is used to solve problems, these problems and their solutions are not unique but may be found and adapted in various contexts and situations (Aldin & de Cesare, 2011).

According to Riehle and Zlighthoven (1996), patterns have two distinct application areas. They are used to create new models in forward engineering, whereas in reverse engineering existing processes can be analysed regarding the existence of the predefined patterns (Forster & Engels, 2003; Becker & Klingner, 2014). Business process patterns in forward engineering are a way of increasing the efficiency and effectiveness of process modelling by reusing existing business functions (Thom et al, 2009; Becker & Klingner, 2014). Therefore reuse of existing process patterns can simplify and speed up the construction of a new model (Aldin & de Cesare, 2011). When a new process model needs to be designed, the user

may start by first identifying the patterns that best satisfy the requirements of his or her context of use. If a process pattern is an excellent fit to a set of requirements it can be used without change. Then the user follows the pattern exactly to design a new process model. Even when a standard pattern does not exactly match the user's requirements, the user may consider changing their requirements and adapt to the pattern.

For their reuse advantages, process patterns are increasingly attracting the interest of both researchers and vendors (Aldin & de Cesare, 2011). Today there exist several proposals for process patterns. One of these proposals is a set of patterns for business process transactions by UN/CEFACT Modelling Methodology (UMM) (UN/CEFACT, 2011). They include transaction and collaboration patterns. *Business transaction patterns* are aimed at providing an established semantics of recurrently occurring business interactions, whereas *business collaboration patterns* defines the orchestrations of activities between partners by defining a set of *business transaction patterns* and/or more collaboration patterns (Jayaweera, 2004). Each pattern, in the UMM, carries a set of default quality of service attributes, such as temporal and security requirements (Wohed et al., 2007).

Workflow patterns are another form of business process patterns that are widely used for modelling business processes (van der Aalst & ter Hofstede, 1999). The Workflow Patterns Initiatives introduces a number of patterns with the aim of outlining the basic requirements that arise during business process modelling on a recurring basis and describe them in an imperative way (van der Aalst & ter Hofstede, 1999; van der Aalst, 2011). Workflow patterns are the atomic process blocks for solving common control-flow data or resource problems that can be used for implementing business requirements in process models (van der Aalst et al, 2003a; Wohed et al., 2007). The *control-flow patterns* capture aspects related to control-flow dependencies between various activities of a business process (Vom Brocke & Rosemann, 2010). The *data patterns* capture the various ways in which data is represented and utilized in workflows (Natschläger & Geist, 2013). Whereas the *resource patterns* capture the various ways in which resources are represented and utilized in workflows (Natschläger & Geist, 2013).

In addition to the workflow patterns there is the Service Interaction Patterns (Barros et al, 2005a; Barros et al, 2005b). While the focus of workflow patterns is on intra-organizational processes, Service Interactions Patterns models the basic service-based interaction blocks used in the design of inter-organizational processes (Barros & Borger, 2005; Barros et al., 2005b).

The outlined patterns can be used to support the design of process models to achieve certain goals, such as temporal goals, risk mitigation goals and others (Wohed et al., 2007). Therefore, business process patterns can serve as the process knowledge that can be stored, as independent reusable objects or together with process models, in the repository to support reuse of process models. Supply Chain Operation Reference Model (SCOR) is good example of a repository that provides standards process patterns for describing supply

chains (Supply Chain Council, 2003). The process patterns provided by SCOR are based the five basic processes: plan, source, make, deliver and return (Glushko & McGrath, 2008). They are organized as conceptual models whose two lower levels of detail refine the basic five processes to describe supply chain models for different industries (Glushko & McGrath, 2008).

In fact, there exists several other business process patterns apart from the ones we have mentioned above, but there is no single pattern that is generally accepted for business process modelling. Since the process patterns are different they require different approaches to define, capture, store and represent them. Therefore the design of the repository to support reuse of process models may vary based on patterns that are to be shared.

2.4.1.2 Process Model Fragments

Process model fragments are reusable process artefacts that are gaining momentum in the field of BPM (Gschwind et al, 2008; Schumm et al, 2011b). There are several different definitions and interpretations of a process fragment. Often, a process fragment is understood to be a group of connected process elements that bear high potential reusability in modelling new business processes (Ma & Leymann, 2008; Schumm et al, 2011a). Process elements are the modelling notation constructs such as activities, control flows, data flows, and routing gateways (Curtis et al, 1992; Ma & Leymann, 2008; Schumm et al, 2011a). A process fragment must at least consist one activity and there should be a way to complete it to a complete process model (Schumm et al, 2011a). It may comprise several incoming and outgoing control flows for integration, with other fragments or into a process (Schumm et al, 2011a). The connectivity of process elements is an essential attribute that simplifies the determination of its control flow included in the resulting process fragment (Ma & Leymann, 2008).

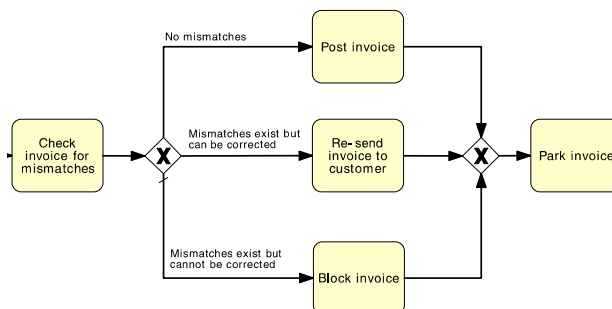


Figure 7: Process model fragment for invoice checking.

Figure 7 is an example process model fragment for invoice checking. First, once an invoice is received it must be checked for mismatches. Three options may result following this check: (i) there are mismatches in which case the invoice is posted, (ii) there are mismatches but they can be

corrected, in which case the invoice is resent to the customer, (iii) there are mismatches but these cannot be corrected, in which case the invoice is blocked. Another interested party can then reuse this fragment.

Process fragment concepts are very similar to workflow patterns, however there is a significant difference between the two concepts. Workflow patterns describe different modelling notation constructs of workflows on an abstract level, whereas process fragment represents concrete functionality (Schumm et al, 2011a).

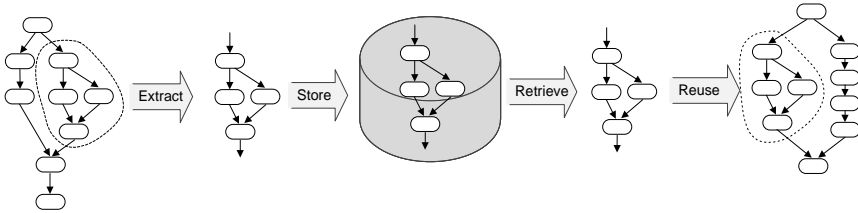


Figure 8: Process model fragment for reuse of process logic (Schumm et al, 2011a).

It has well been established that reuse of process model fragments in process modelling improves the quality of models and increases the efficiency of modelling (Markovic & Pereira, 2008a; Ma & Leymann, 2008). The process fragments can be stored in the repository and used in modelling other processes by retrieving a process fragment that realizes particular process logic (Ma & Leymann, 2008; Schumm et al, 2011b). A good example of a repository that provides process fragment for reuse is Fragmento as proposed by Schumm et al (2011b). Fragmento manages BPEL process fragments and other process-related artefacts. Apart from providing basic repository functionalities, Fragmento provides some advanced functionalities such as process fragmentation techniques. As shown in Figure 8, in order to reuse process fragment, at first a process logic fragment needs to be either extracted from a process model or designed from scratch (Schumm et al, 2011a), before it is stored in the repository for reuse. Therefore repository design must take into consideration how process logic can be extracted from complete process models. Thus the decision of providing process fragments as reusable process knowledge may influence the repository design.

2.4.1.3 Business Process Ontologies

An ontology is an explicit specification of conceptualization (Gruber, 1993; Noy & McGuinness, 2001; Bullinger, 2008). A specification of conceptualization is defined as the objects, concepts, other entities and the relationship between them, which are assumed to exist in some area of interest (Gruber, 1995). Ontologies have widely been recognized as a means to systematically design business processes (Lin & Strasunskas, 2005; Hepp et al, 2005; Hepp & Roman, 2007). Most of the work on adopting ontology to processes modelling focuses on creating process representations aimed at

facilitating the automated processing of models (Aldin, 2010, Aldin & de Cesare, 2011). These works are geared towards a particular modelling language – they mainly deal with semantic annotation of process models represented with the help of a particular modelling language (Aldin, 2010, Aldin & de Cesare, 2011).

Several works on the application of ontologies in process modelling exist. These works include the SUPER ontology by the Semantics Utilized for Process Management within and between Enterprises (SUPER) project (SUPER, 2004; Markov & Kowalkiewicz, 2008). In this work Semantic Web standards have been used to develop business process management (BPM) approaches (SUPER, 2004; SUPER, 2007). Focusing on bridging the gap between business and IT using Semantic Web technologies, the project enables semi-automation of the BPM life cycle (Nitzsche et al., 2007; Aldin & de Cesare, 2011). The SUPER ontology provides an extensive conceptualization of the BPM domain – covering from modelling of business processes to defining business strategies (SUPER, 2007; Pedrinaci et al, 2008). To achieve this, a set of constructs is provided to produce semantic business process models by applying ontologies to business processes (Aldin, 2010; Aldin & de Cesare, 2011). In order to formalize process knowledge representation, several ontologies for different modelling languages, such as BPMN, EPC, BPEL, and Petri Nets have been proposed.

Brockmans et al (2006) is another proposal for the application of ontologies in business process modelling is that proposed. In their work they propose a semantic alignment approach for process modelling using Petri nets. A representation of Petri nets is provided in an ontology language such as the Web Ontology Language as a way to semantically enrich the process models (Aldin, 2010).

Another application of ontologies to business process modelling is that by Thomas and Fellmann (2006), who propose annotating EPCs with semantics (sEPC). In this proposal four instances of ontologies are included – business ontology, business process concepts, the sEPC model and the underlying EPC model (Lautenbacher & Bauer, 2007).

For the repository to support the use of ontologies as the foundation of business process modelling it requires mechanisms and interfaces or plug-ins to annotate process models with intended ontologies. Consequently the use of business process ontologies may influence the design of the process model repository.

2.4.2 Process Representation and Understanding

As shown in Figure 6, before a new model is composed from existing models, users need to understand it and then modify or adapt the model for the new business requirements. Process model understandability is defined as the degree to which the reader of that model can easily understand information contained in the process model (Mendling & Strembeck, 2008; Reijers & Mendling, 2011; Mendling et al, 2012).

Factors that influence the process model understandability can be categorized as personal factors and model factors (Reijers & Mendling, 2011; Mendling et al, 2012). Personal factors relate to the reader of the model, and include user experience and education as related to process modelling, whereas model factors relate to the process model itself (Reijers & Mendling, 2011). Process model factors that influence the comprehension of the syntactical content include process model complexity (Mendling & Strembeck, 2008; Lucrédio et al, 2008; Mendling et al., 2012), modelling notation (Agarwal et al, 1999; Sarshar & Loos, 2005), visual representation (Moher et al, 1993; Purchase, 1997; Reijers et al, 2011), model purpose (Reijers & Mendling, 2011) and problem domain (Lakhotia, 1993).

Some of the above factors are related to the intended use of the process model. In particular, the problem domain and the model purpose are typically an intentional process design decision. Other factors such as user experience and education cannot be addressed by the repository design. With respect to process model complexity, several metrics have been proposed for measuring the complexity of process models (La Rosa et al, 2011b; Ghani et al, 2008; Cardoso et al, 2006). The metrics include the size of the model, which simply counts the number of activities in a business process (Ghani et al, 2008; Laue & Mendling, 2010). Another metric is the structure of the model, which depends on the nesting depth contained in the model – a higher nesting value indicates a more complex model (Cardoso et al, 2006). While a considerable amount of literature is devoted to this topic, there is no straightforward solution to addressing the complexity of process models (La Rosa et al, 2011b). However, abstraction of process models has widely been considered as a way to facilitate the user's conception of the contents of the complex process model (Smirnov et al, 2012; Polyvyanyy et al, 2010). This requires the repository to provide a mechanism to abstract complex process models to specific use cases that can easily be understood by the user.

Process representation, as it relates to modelling notation and visual representation factors, is one of the main factors that must be addressed by the repository to improve process model comprehension and reuse (Mendling & Strembeck, 2008; Schrepfer et al, 2009). A representation is a language (textual, graphical or other) used to describe a process model (Eid-Sabbagh et al, 2012; Elias & Shahzad, 2010). As stated in Section 2.2, there are several modelling notations, which are used to represent business processes either in graphical or textual form. These modelling notations have different elements and control structures (Lu & Sadiq, 2007), therefore the specification of a business process varies from one notation to another (Shahzad et al., 2010). Some repositories, such as SAP's Business Map (SAP AG, 2007), provide graphic-based descriptions, while others, such as MIT's Process Handbook (MIT process handbook project, 2001), describe processes only in a textual form.

Process representation also influences the ease with which a process model can be modified to meet user requirements. One of the primary benefits of abstracting a business process using process models is that they

allow automated enactment by using PAIS (Dumas et al, 2005; Indulska et al, 2009). PAIS requires the process logic be expressed in terms of executable process models (Dumas et al, 2005). This implies that process models in the repository must be represented in a format that can easily be transformed into executable solutions. Existing repositories describe process models in textual or high-level forms, which cannot be directly executed by using PAIS or be easily transformed into an executable form (Wohed et al., 2007).

In order to increase process model reuse the representation of process models in the repository should be addressed to facilitate comprehension and easy transformation of process models. As previously mentioned, there are several modelling notations in which a process model can be represented (Elias & Shahzad, 2010). For designing a publicly open repository the issue of representation of process models does not have simplistic solutions and needs to be addressed separately. Possible process representation options for such a repository include the use of one standard modelling notation, multiple standard modelling notations and a compatible modelling notation (Elias & Shahzad, 2010; La Rosa et al, 2011a).

2.4.2.1 One Standard Modelling Notation

In this option one standard modelling notation, such as BPMN, YAWL or EPC, is selected and supported in the repository. This will simplify the design of the repository, however it will affect the possibility of sharing process models described using another representation. Companies and researchers have already created a large number of process models using different modelling notations. Therefore restricting the repository to one modelling notation will limit contributors to sharing their existing models, consequently affecting the goal of achieving a critical mass of process models for reuse.

2.4.2.2 Multiple Standard Modelling Notations

In this option the repository supports multiple modelling notations. This will allow process models to be stored and retrieved in their respective modelling notations. This will avoid losing process models represented in another format and hence achieving a critical mass of process models for reuse. However, it will make the repository design and implementation more difficult than supporting one standard notation. A possible solution is that proposed by (Peters, 2007), which requires multiple modelling language definitions (stencil sets) to be defined and maintained in the repository.

2.4.2.3 One Compatible Modelling Notation

In this option the repository supports one modelling notation, which is compatible with other modelling notations. This requires the repository to provide a mechanism to convert process models into a single format. The conversion can be applied when models are populated in the repository. A good example is the work by La Rosa et al. (2011a), which provides a

canonical format – a format compatible with different modelling tools that can be used to address the challenge of supporting multiple modelling languages. The main disadvantage with this option is that it leads to information loss during conversion from a specific notation to the common format. This is because the semantics of the original model will not be preserved during the model conversion. While such semantic loss may be less challenging for a repository that is aimed at providing basic understanding and fundamentals of process models for reuse, it may not be acceptable for certain applications (Shahzad et al., 2009).

In general, the choice of how process models should be represented often influences the design of the repository. The repository should be designed to ensure that process models are represented in a format that is easy to understand and adopt or modify for reuse. However, the choice should take into consideration the need to achieve a critical mass of process models.

2.4.3 Repository Structuring and Retrieval

One of the main issues facing process model repositories is managing a large collection of process models efficiently and allowing fast allocation and retrieval (Elias & Johannesson, 2012b; Yan et al, 2010). A major goal of reuse is, of course, to “find the artefacts faster than the time it takes to develop them”. Therefore, providing a mechanism to enable faster retrieval of the relevant process models is one of the main repository design considerations. Repository structuring is a fundamental factor in achieving good retrieval results. Despite the fact that some retrieval algorithms can offer sufficient effectiveness with marginal structuring and indexing efforts, a poorly structured repository with insufficiently indexed contents will not have a good retrieval performance irrespective of the retrieval algorithm (Shiva & Shala, 2008).

The structuring of the repository is commonly achieved through the use of classification schemes (Ali & Du, 2004; Prieto-Diaz & Freeman, 1987). Classification schemes have long been used to give structure to a large body of information (Ali & Du, 2004; Prieto-Diaz & Freeman, 1987). A well-established classification scheme can improve the understanding and retrieval of the repository content, thus making the content more beneficial (Ali & Du, 2004; Prieto-Diaz & Freeman, 1987). It provides a means by which artefacts can efficiently and effectively be stored and retrieved. Properties of classification schemes allow the representation of stored artefacts and relationships to structures that reflect knowledge of the domain being classified (Ali & Du, 2004).

In this thesis, the issue of structuring the repository to improve the performance of the retrieval system is one of the main issues we have tried to address. This is related to the second goal of the research, which is *to develop and evaluate a process semantic annotation model for semantically*

annotating process models in the repository. We have developed and evaluated a context-based process semantic annotation model for semantically annotating business processes in the repository (Elias & Johannesson, 2013; Elias et al, 2010). The model is based on existing process classification schemes and business frameworks. The annotation model is discussed in Chapter 5. As a proof concept the annotation model was implemented in a repository prototype, which was used for evaluation of the model in the controlled experiments (Elias & Johannesson, 2012a). The evaluation of the model is presented in Chapter 7.

As part of the repository structuring we have developed a business process relationship meta-model for identifying and defining the relationship between process models in the repository. This is related to the third goal of our research, which is to *develop and evaluate a model for identifying and defining the relationship between process models in the repository.* While the semantic annotation model is aimed at facilitating both searching and navigation, the meta-model relationship extends the navigation to processes that are related. The meta-model is based on existing process relationships (Kurniawan et al, 2012; Malone et al, 2003) and process-assets and asset-processes archetypes we have developed as a way to find all business processes in an enterprise (Bider et al, 2012; Elias et al, 2014). The process relationship meta-model and the evaluation are described in Chapter 8.

2.5 Summary and Discussion

This chapter has sketched the context of this research work. We have introduced the theoretical and technical settings relevant to our research topic, such as business process management, business process modelling, process model reuse and process model repository. The main points can be summarized as follows.

- Business process management (BPM) is an important discipline that focuses on improving corporate performance by managing and optimizing a company's business processes.
- Process modelling is a critical component for successful implementation of BPM.
- Modelling business processes of modern enterprises is complex, time-consuming and error-prone and as a result it requires extensive efforts to produce quality models.
- An alternative for reducing the effort to modelling business processes from scratch is redesigning process models from existing ones.
- A process model repository that is properly populated with a critical mass of process models is a step towards supporting reuse of process models.
- Providing a mechanism to enable faster retrieval of the relevant process models is one of the main repository design considerations.

- Restructuring of process model repositories is key to providing an effective repository retrieval system.
- Process classification is an approach to structuring process model repositories.
- A semantic annotation model that includes process classification and meta information will facilitate searching, navigation and understanding of process models.

Having established the key areas of our research in this thesis, we will continue with the research design of the thesis. A discussion of related work will be presented in the respective section of our contributions – the requirements for the repository, the semantic annotation model, the process relationship meta-model and the repository architecture.

3 Research Design

Research design defines the scope and activities within a research project (Morrow & Brown, 1994; Denscombe, 2010). It is a structured sequence of steps within the research project that encompasses the discussion of the initial research problem to the final conclusion.

This chapter presents the research design of this thesis. We begin by providing the methodological foundation of the thesis followed by the research process and the research methods used in this thesis.

3.1 Research Paradigm

The field of information systems deals with technological artefacts (computerized systems) in non-technological settings (human organizations). The two different sides normally call for different research methods. As a result, research in the information system discipline is mainly characterized by two paradigms – behavioural science and design science (Hevner et al, 2004). Advocates of the behavioural science paradigm view the information systems discipline as a social science, while those from the design science paradigm view it as engineering (Hevner & March, 2003).

The behavioural science paradigm is rooted in natural science research methods. It looks for the development and verification of theories that explain organizational and human phenomena around information systems' life cycle. Therefore the goal of behavioural science is to produce knowledge by answering a knowledge question (Hevner & March, 2003). The solutions to such questions are propositions that are claimed to be true.

The design science paradigm is rooted in engineering and the sciences of the artificial (Simon, 1981) and is basically a problem-solving paradigm. According to Denning (1997), design science research looks for the creation of innovations that define the ideas, practices, technical capabilities and products necessary to accomplish the development and use of information systems. Therefore the goal of design science is not only to create knowledge, but also to solve practical problems.

Design science creates and evaluates information technology artefacts aimed at solving identified organizational problems (Hevner et al, 2004). While several attempts to define the IT artefact exists, a well-accepted suggestion is that proposed by March and Smith (1995). They differentiate between four types of IT artefact: constructs, models, methods and instantiations.

- Constructs are the conceptual vocabulary of a domain (Schon, 1984). They provide a language that can be used to define and communicate problems and solutions.
- Models are a set of statements or propositions used to express the relationships between constructs. They represent a real-world situation – the design problem and its solution space (Simon, 1981).
- Methods are sets of steps that can be used to solve the designated problem. They provide guidance on how to solve problems – searching the solution space.
- Instantiation is the operationalization of constructs, models and methods.

According to March and Smith (1995), and Joahannesson and Perjons (2012), design science research must eventually lead to one of these artefacts. Therefore design science research is responsible for methodically guiding the development of an artefact.

The main problem that this research is addressing is *how to design an open and language-independent process model repository with an efficient retrieval system to support reuse of process models*. Since this is a practical problem and the solution to this problem includes the development of several artefacts, we have chosen to follow a design science research paradigm. The artefacts that we intend to construct include a semantic annotation model for semantically annotating process models in the repository to facilitate efficient retrieval of process models, and a business process relationship meta-model for identifying and defining relationships between process models, and the architecture of the process model repository. In addition, the research work may include the instantiation of the proposed architecture, models and methods by implementing a working prototype of the repository.

The construction of an artefact does not directly provide any empirical data, however the knowledge that is required in the design process comes from empirical and theoretical sources. Therefore, while the primary methodical approach that is applied in this thesis is design science research, several complementary research strategies and methods have been employed to prepare and evaluate the design process and to collect the necessary data. In addition, during problem identification additional research strategies and methods are needed to gain a sufficient empirical basis. The complementary research strategies employed in this thesis include survey, case study and laboratory experiments (Hevner et al, 2004; Peffers et al, 2012), whereas the research methods employed include interview, questionnaire and document analysis.

A detailed description of the research process undertaken in carrying out this research is provided in Section 3.2. Below we describe the activities of a typical design science research process that we have chosen as a template.

3.1.1 Design Science

Activities of design science research projects naturally group into several subprocesses or phases. Peffers et al (2007) divide these activities into six subprocesses, described below and shown in Figure 9.

Firstly, the research problem that is sought be addressed is identified and motivated. Secondly, the objective of the solution is inferred from the problem specification. The third activity is the design and development, which creates the actual artefact to address the problem. As the fourth activity, the efficacy of the artefact is demonstrated to determine whether it solves the problem. The fifth subprocess is when the produced artefact is evaluated to measure the extent to which it meets the solution requirements in solving the problem. Finally, as the sixth subprocess, the research results are communicated to researchers and practitioners. The communicated results should range from the problem and its importance to the artefact and its utility and the rigour of its design.

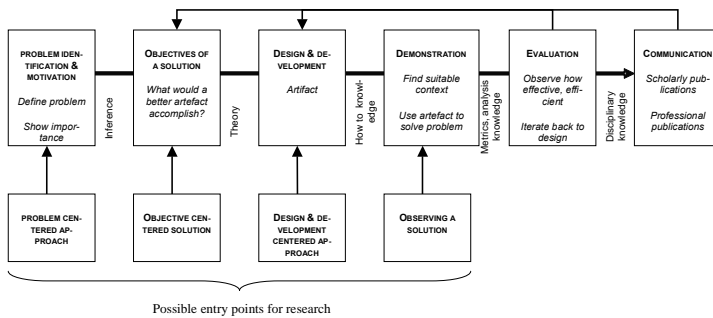


Figure 9: A design science research process (Peffers et al, 2007).

3.1.1.1 A Template of a Design Science Research Process (TDSRP)

In this section we present a template of a design science research process (Joahannesson & Perjons, 2012) that we have adopted to describe the research process of this thesis. The TDSRP is an amendment of the design science research process proposed by Peffers et al (2007). Similarly the TDSRP shown in Figure 10 decomposes the research process into six activities, namely (Joahannesson & Perjons, 2012): explicate problem, outline artefact and define requirements, design and develop artefact, demonstrate artefact, evaluate artefact, and communicate artefact knowledge. In this template a research process takes the initial problem to be solved as the input, and produces the artefact and the artefact knowledge as the outputs (Joahannesson & Perjons, 2012). In addition, the process is administered by research methods as controls and uses knowledge foundation as mechanisms. Below we describe the activities.

Explicate problem. This is the first activity of design science research process. In this activity the initial problem is precisely formulated and motivated, and its underlying clause is investigated. It answers the question

“what is the problem experienced by some stakeholders and why is it important?” (Joahannesson & Perjons, 2012). This activity takes as input the initial problem that is unclear and produces an explicated problem that is clearly defined, motivated and contextualized.

Outline artefact and define requirements. This is the second activity of design science research process. In this activity, an artefact that can address the explicated problem is identified and outlined, and its requirements are defined. It answers the question “what artefact can be a solution to the explicated problem and which requirements of this artefact are important to the stakeholders?” (Joahannesson & Perjons, 2012). This activity takes the explicated problem as input and produces an outline of the artefact and its requirements as the output.

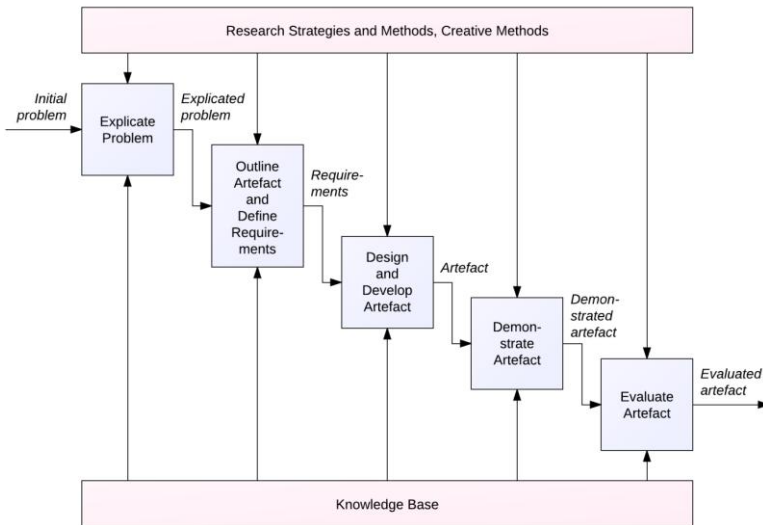


Figure 10: A template of design science research process (Joahannesson & Perjons, 2012).

Design and develop artefact. This is the third activity of design science research process. In this activity the artefact that fulfils the defined requirements and addresses the explicated problem is developed. It takes as the input the artefact outline and its requirements and produces the artefact that fulfils the requirements and associated knowledge.

Demonstrate artefact. This is the fourth activity of design science research process. In this activity the use of the artefact is shown in one or a number of cases to prove its feasibility. It answers the question “how can the developed artefact be used to address the explicated problem in one case?” This activity takes artefact as input and produces the demonstrated artefact with information on the working of the artefact in one case.

Evaluate the artefact. This is the final activity of design science research process. In this activity the artefact is evaluated to determine how well it is able to solve the explicated problem and the extent to which it meets the requirements. It answers the question “how well does the artefact solve the

explicated problem and fulfil the defined requirements?” (Joahannesson & Perjons, 2012). This activity takes an artefact as input and produces an evaluated artefact with information about how well the artefact works and why.

3.2 The Research Process

The research process specifies the activities within a research project and defines how they interact. It relates the research goals and the research strategies and methods to particular phases of the project.

In this section, we describe the research processes behind this thesis. The research processes are described below and their corresponding graphs presented in Figures 11, 12 and 13.

3.2.1 The Research Process for the Overall Research Goal

The first research process is the process behind the overall research goal of the thesis. It aimed: “to design an open and language-independent process model repository with an efficient retrieval system to support reuse of process models”. Below we describe the research process as depicted in Figure 11.

Explicate the problem. The first activity of this research process was to explicate the problem. In this research, the problem is drawn from a literature review on process model reuse. Despite their benefits, the existing process model repositories suffer from several shortcomings that affect their usability in practice. Therefore this research process is addressing the following explicated problem: “*how to design an open and language-independent process model repository with an efficient retrieval system to support reuse of process models?*”

Outline artefact and define requirements. The second activity of this research process was to outline the artefact and define the requirements. The artefact designed in this research process is the architecture of a process model repository. To better understand and address issues and challenges affecting existing repositories it is vital to investigate the requirements for building such repositories. It is during this activity that the first goal of this thesis was achieved, that is “*to establish the requirements for the process model repository*”. The requirements have been elicited from two different sources: stakeholders, and literature survey, i.e. documents and existing systems.

Requirements elicitation from stakeholders. To understand the requirements for a process model repository, we chose an *exploratory approach*. This is because at this point the problem that limited reuse of process models was not well understood. In addition, very little research on process model repositories existed. The study was separated into two phases of (1) generation of preliminary requirements and (2) a confirmatory study,

which overall involved two categories of stakeholders, i.e. practitioners and researchers. In the first phase we chose an interview since the method focuses on information depth and enabled investigation of diverse requirements (Denscombe, 2010). We used a one-to-one semi-structured interviews based on a predefined interview instrument (Appendix A.2). During the interview, audio data were recorded as audio, which were then transcribed for analysis. The transcript data were analysed using a thematic analysis technique (Ritchie & Spencer, 2002). The analysis of the five responses from the interviews produced requirement propositions that expressed suggestions for requirements for a process model repository (Elias & Johannesson, 2012). In order to validate these propositions, we conducted a *confirmatory study* with a larger set of 30 participants (researchers and practitioners) who participated in the second Working Conference on the Practice of Enterprise Modelling (POEM 2009) (Zdravkovic, 2009). A questionnaire was used for validating the requirement propositions. The questionnaire included additional column to allow for more comments/ideas/suggestions from the respondents (see Appendix A.3). The analysis of the 25 responses produced requirements for a process model repository.

Requirement elicitation from literature. The requirements from the literature (documents, existing process model repositories, etc.) were elicited through a systematic review approach (Kitchenham, 2004; Kitchenham & Charters, 2007). The process consisted of three major steps: planning, conduction and requirement establishment. In the planning step we defined the research objectives and the review protocol. In the conduction step we identified, selected and evaluated the primary studies according to the inclusion and exclusion criteria established in the defined protocol. This was based on publications in academic as well as trade journals and conferences. To identify the primary studies we searched relevant journals by using Google Scholar. For each selected study, data were extracted and synthesized. Finally, in the requirements establishment step the requirements were established through brainstorming in the discussion group.

Design and develop artefact. The third activity of this research process was design and develop the architecture of the process model repository. It was during this activity that the fourth goal of this thesis was achieved. The method used to design the architecture was a synthesis-based architecture design (SYNBAD) approach (Tekinerdogan & Aksit, 2002). SYNBAD was chosen because the scoping of the architecture is not only based on the stakeholders' perspective, but also on a systematic problem-solving perspective (Tekinerdogan & Aksit, 2002). The design of the architecture consisted of four iterative subactivities: technical problem analysis, solution domain analysis, create architecture and describe the architecture.

The first subactivity was *technical problem analysis*. In this subactivity, the requirements specified in the previous activity were generalized and then mapped to technical problems. Technical problems were then decomposed into subproblems, which could be solved individually. Before moving to

solve the subproblems, however, prioritization was carried out to determine which subproblem needed to be solved first.

The second subactivity was *solution domain analysis*. In this subactivity, we identified and prioritized the solution domains for every subproblem defined in the technical problem analysis. This was done through discussion and systematic literature search using Google Scholar. Then the knowledge sources, for each solution domain, were defined and prioritized. After studying and analysing the solution domain knowledge, the fundamental architectural concepts were extracted. The concepts were then structured to create the architecture flow chart using relations that are derived from the solution domains.

Prior to creating the final architecture, in the next subactivity, two conceptual models that represented architecture information models were developed and evaluated. These were: (1) a semantic annotation model for semantically annotating process models in the repository to facilitate understanding, searching and navigating the repository, and (2) a business process relationship meta-model for identifying and defining the relationship between business processes in the repository to facilitate navigating the repository. The construction of the two models marks the achievement of the second and third research goals, which are the main contributions of this research work. The detailed research processes for developing the two models are presented in Sections 3.2.2 and 3.2.3. Below we describe the research process for the design and development of the architecture.

The third subactivity was to *create the architecture*. The concepts represented in a conceptual structure were then used to create the architecture by instantiating the Service Oriented Architecture (SOA) reference architecture as an architectural style (Erl, 2005). SOA was chosen because it allows for increased scalability, flexibility and integration in comparison to other architectural styles (Erl, 2005).

The fourth subactivity was to *describe the architecture*. The architecture is described based on the ISO/IEC/IEEE 40120 standard (ISO/IEC/(IEEE), 2011), as it offers different viewpoints for describing the system architectures.

Evaluate the artefact. The fourth activity of this research process was to evaluate the architecture as the artefact. To evaluate the proposed architecture, we used the informed argument (Hevner et al, 2004; Peffers et al, 2012) evaluation method. This method was appropriate as it allowed the use of requirements specification as the knowledge base from which the utility of the proposed architecture was examined. We performed a systematic analysis on how the architecture fulfils each of the established requirements. In addition, a prototype that implemented some of the components of the architecture serves as a proof of concept to show the feasibility of the architecture implementation.

The theoretical foundations used in the research process were business process management, requirements engineering, software and systems engineering, reusable object repository theory, software component reuse

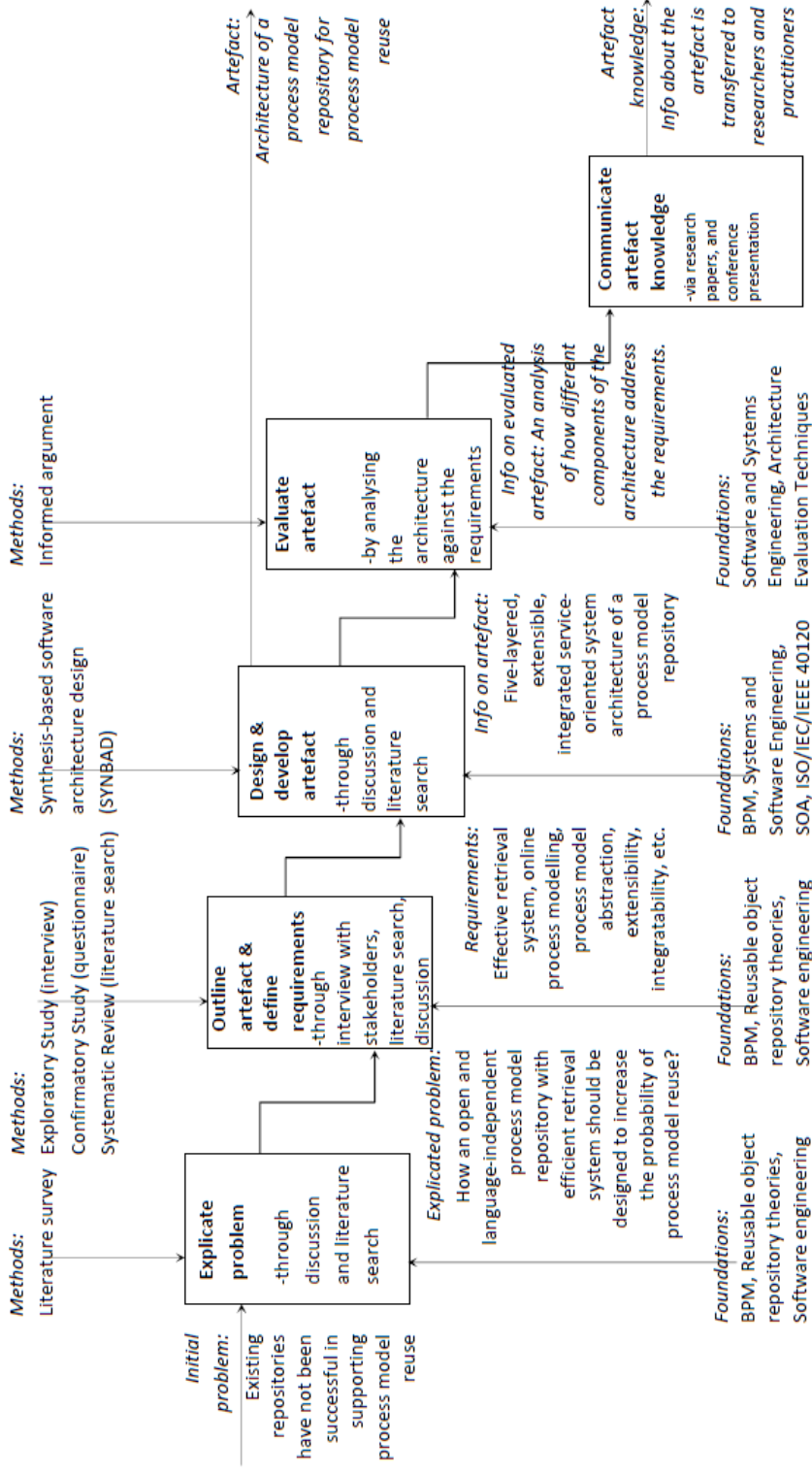


Figure 11: Research process for the overall research goal.

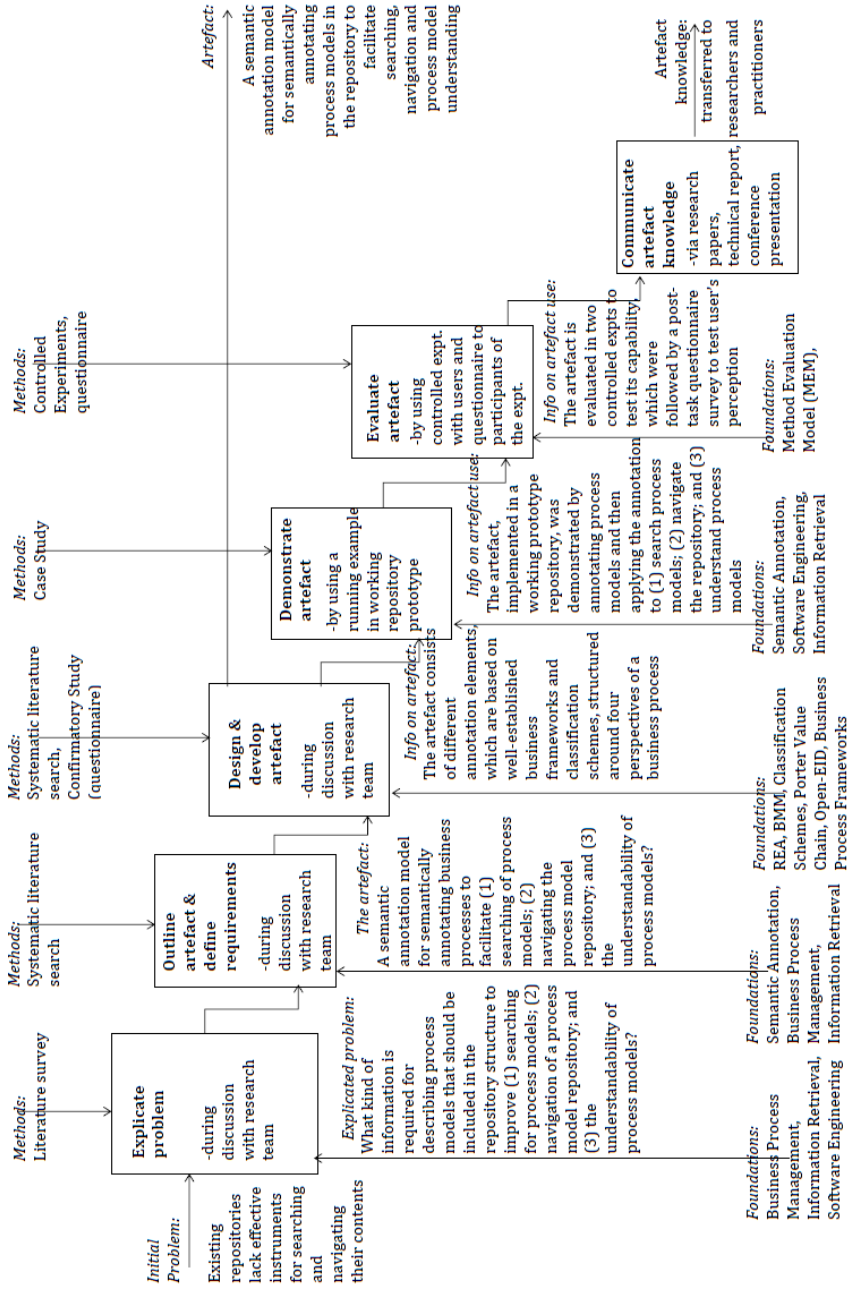


Figure 12: Research process for achieving goal 2 of the thesis.

knowledge, service oriented-architecture (SOA), and software architecture evaluation methods.

3.2.2 The Research Process for Goal 2

The second research process is the process behind the second goal of the thesis. It aimed: “to develop and evaluate a semantic annotation model for semantically annotating process models in the repository”. Below we describe the research process as depicted in Figure 12.

Explicate problem. The first activity of this research process was to explicate the problem. In this research, the problem is drawn from a literature review and a survey (Elias & Johannesson, 2012b) of existing process model repositories, carried out to identify the challenges that limited their use. One of the main findings is that the existing repositories lack effective instruments for searching and navigating their content. In order to use, users must be able to find process models with less effort (efficiently) and that meet their business needs (effectively). Therefore this research process is addressing the following explicated problem: what kind of information is required for describing process models that should be included in the repository structure to improve (1) searching for process models; (2) navigation of a process model repository; and (3) the understandability of process models?

Outline artefact and define requirements. The second activity of this research process was to outline the artefact and define the requirements. The artefact designed in this research process is a semantic annotation model for semantically annotating process models stored in the repository to facilitate searching, navigation and process model understanding. In accordance with design science research (Peffer et al, 2007; Joahannesson & Perjons, 2012), a set of requirements for the annotation model were established to guide the design and development activity. The requirements were elicited through a systematic literature search – the review of related work (Kitchenham & Charters, 2007). In addition, the lesson learnt from the overall requirements of the repository (in the first research process) and review of existing repositories provided more input to the establishment of the annotation model requirements.

Design and develop artefact. The third activity of this research process was to design and develop the semantic annotation model. The specified requirements were used as input to the development of the annotation model. The design and development of the annotation model consisted of three steps. In the first step we identified potential annotation elements through literature search and document studies. In the second step, the identified elements were validated in a “confirmatory study” using a questionnaire. To accommodate new annotation elements from the respondents, the questionnaire was designed (see Appendix A.4) to allow them to provide comments on each potential element or suggestions for additional potential

annotation elements. Finally, the validated annotation elements and the feedback from participants were used to construct the annotation model through brainstorming in group discussions.

Demonstrate artefact. The fourth activity of this research process was to demonstrate the artefact. The application of the annotation model was demonstrated using a running example as a “case study”. A repository prototype that implements the annotation model was used to annotate and store process models in the repository. This was followed by a demonstration of how the annotation-based search (ABS) mechanism is used to find relevant process models, how to navigate the repository based on the annotation and how the annotation helps a user to understand process models. The implementation of the annotation model in the repository prototype and running example served as a proof of concept showing that the model could be used as intended.

Evaluate artefact. The last activity of this research process was to evaluate the designed artefact. An *ex post* strategy was chosen, i.e. the artefact was deployed for evaluation (Joahannesson & Perjons, 2012). We adopted a method evaluation model (Moody, 2003) as a framework to evaluate the annotation model. The annotation model has been evaluated in two different studies using controlled “experiments” to test its performance on model retrieval. The experiment was chosen since it is the most appropriate evaluation strategy for evaluating the performance of models and methods in design science research (Peffer et al, 2012). In both studies, user perceptions of the annotation model are evaluated using post-task “survey questionnaires”, which followed experimental tasks.

The theoretical foundations used in this research process were information retrieval knowledge, business process management, semantic annotation concepts, business frameworks (i.e. REA, BMM, Open-EDI, Porter Value Chain), and the method evaluation model.

3.2.3 The Research Process for Goal 3

The third research process is the process behind the third goal of the thesis. It aimed: “to develop and evaluate a model for identifying and defining the relationship between business processes”. Below we describe the research process as depicted in Figure 13.

Explicate problem. The first activity of this research process was to explicate the problem. One of the issues in the design of the process model repository is the need to identify and define the relationship between process models in the repository. Therefore this research process is addressing the following explicated problem: “*what kind of relationships between business processes should be defined to improve navigation and usage of process models in the repository?*”

Outline artefact and define requirements. The second activity of this research process was to outline the artefact and define the requirements. The artefact designed in this research process is a business process relationship meta-model for identifying and defining the relationship between business processes in the repository. In accordance with design science research (Peffers et al, 2007; Joahannesson & Perjons, 2012), a set of requirements for the relationship meta-model were defined to guide the development of the model. The requirements for the meta-model were elicited through a systematic literature search – the review of related work.

Design and develop artefact. The third activity of the research process was to design and develop the business process relationship meta-model. The design of the relationship meta-model has been achieved through a three-phase process.

Phase 1: In the first phase, we identified potential process relationship concepts that can be used in the repository. Two sources were used for identifying the relationships between business processes. The first source is the existing literature on process model repositories. Here we performed a literature analysis that resulted in a number of relationships. The second source was our own research project that aimed at the development of a procedure to find all processes in an enterprise. Two major types of interconnection between the business processes were identified and defined through group discussion and brainstorming (Bider et al, 2012). They are process-assets and asset-processes archetypes. The process-assets archetypes help to identify all the assets needed to successfully run a business process and the asset-processes archetypes help to find all processes needed to manage the asset. While process-assets and asset-processes archetypes served as a method to find all processes in an enterprise, they also provided relationship semantics that links the business processes of an enterprise.

Phase 2: In the second phase, we validated the process-assets and asset-processes archetypes (Elias et al, 2014). A case study was chosen as a research strategy to evaluate the suitability of the archetypes. This is because our aim was to instantiate the archetypes in a real-world case, in this case the department of computer and systems sciences (DSV) of Stockholm University (SU). The validation consisted of several activities. (i) The first step was an investigation of business processes in the department. The data were collected through one-to-one semi-structured interviews with nine business domain experts (including directors, heads of academic units and operation staff). This allowed us to get an in-depth understanding of how the higher education institution works (Denscombe, 2010). The interviews were based on a predefined interview instrument (Appendix A.6). During the interviews, data were recorded as audio from nine business domain experts, which were then transcribed for analysis. In addition, as a complement to the interviews data were also collected by analysing formal documents provided during the interviews. The analysis of the responses from the interviews and document analysis produced a set of business processes performed at the

department. (ii) In the second step, using the results from the first step we applied the archetypes to build the educational process architecture. This was achieved through brainstorming in discussion groups. (iii) In the third step, the produced process architecture was presented to the business domain experts from the departments. This was followed by evaluation of the process architecture by the business domain experts. A semi-structured interview was chosen for the evaluation. The presentation and interview were also conducted using a one-to-one approach. The analysis of the seven responses from the interviews allowed us to evaluate the archetypes and relationship semantics, which created the process architecture.

Phase 3: In the third phase, we *constructed the meta-model*. This was done first by integrating existing process relationships used for defining and maintaining process model relationships in the repository. Then the existing concepts were integrated with the relationship semantics from process-assets and assets-processes to define the meta-model. The data collected from the case study were used as examples to describe the meta-model.

Evaluate artefact. The last activity of this research process was to evaluate the meta-model. The meta-model was evaluated by analysing how it meets the established requirements. This was achieved through an informed argument (Hevner et al, 2004). The application of the meta-model in the repository has not been tested because it has not been implemented in the repository.

The theoretical foundations used in this research process were enterprise modelling, business process management and the SEQUAL framework.

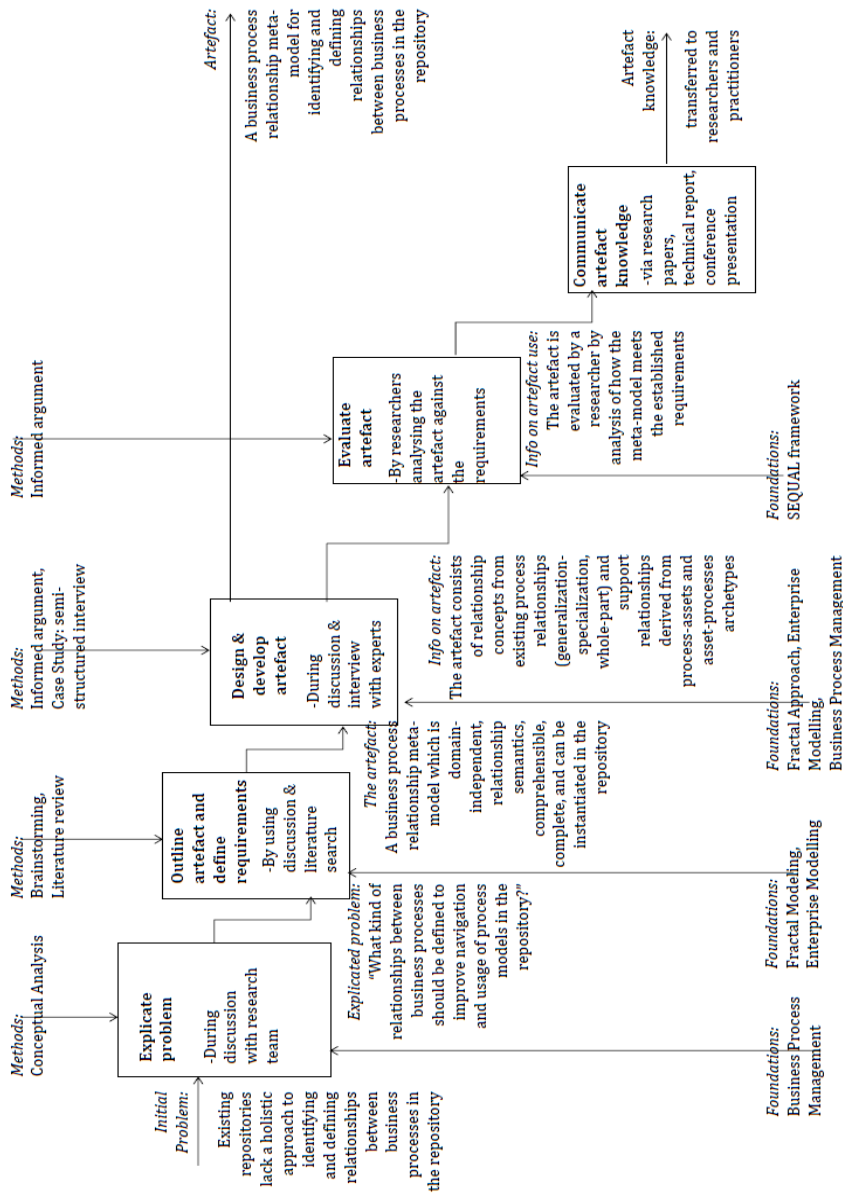


Figure 13: Research process for achieving goal 3 of the thesis.

4 Requirements for a Process Model Repository

4.1 Overview

The chapter investigates requirements for a process model repository. The investigation is intended to seize the most essential requirements that repositories must fulfil in order to increase the probability of process model reuse. We have elicited the requirements from two different sources: stakeholders and existing literature. The chapter presents a step-by-step process that describes how the requirements were elicited, analysed and established from the two sources. We begin by describing the requirements elicitation and analysis process from the stakeholders. This is building on Paper II as listed in section 1.5 of Chapter 1. We show how the requirements were elicited through an exploratory study and then validated through a confirmatory study. A detailed discussion of the analysis of the responses between researchers and practitioners who participated in the confirmatory study is presented. Requirement statements from stakeholders as the results of the analysis are then presented. The chapter also describes the requirement elicitation and analysis process from existing repositories and associated documents from the literature. We show a step-by-step process ranging from panning, conduction and establishment of the requirements statements. A detailed discussion of the search criteria and strategy is provided. We show and discuss how sources were identified and selected, and how the requirements were extracted from the selected sources. We then present the requirement statements from the literature as the results of the analysis. The chapter then presents the specifications and justifications of the requirement statements from both stakeholders and the literature. The chapter also discusses some of the related work. Finally the chapter concludes with some key remarks about the requirements.

4.2 Requirements Elicitation from Stakeholders

Requirements elicitation is the process of seeking, discovering, acquiring, and expounding requirements for the software systems (Kotonya & Sommerville, 1998; Zowghi & Coulin, 2005). Requirements may be obtained from various sources in a variety of formats (Kotonya & Sommerville, 1998; Zowghi & Coulin, 2005). The most obvious source of

requirements for the system is the stakeholders – users and subject matter experts who understand the problems and needs. Another source for eliciting the requirements is the existing systems and processes. Existing documentation about the existing systems can provide useful information about the requirements for the new system that need to be developed and their supporting motivation and significance (Kotonya & Sommerville, 1998; Zowghi & Coulin, 2005). Therefore, in order to establish the requirements for the process model repository we chose to elicit the requirements from two different sources: stakeholders of the system, and existing repositories and documentation about existing repositories.

In this section we present the elicitation and analysis of the requirements for a process model repository from stakeholders. The elicitation and analysis of the requirements from the existing repositories and documentation are discussed in the next section. The process of gathering requirements from stakeholders consisted of two phases as shown in Figure 14: an exploratory study and a confirmatory study.

The first phase was an exploratory study, which aimed at gathering comments, ideas, suggestions, and opinions on requirements for process model repositories from both researchers and practitioners. The study was designed to collect as many suggestions and ideas as possible, encouraging the participants not to restrict their responses. It is therefore, for this reason the exploratory study consisted of an open-ended questionnaire that was used in an oral interview. The collected responses were reformulated into propositions that expressed requirements suggestions for a process model repository. Since the purpose was to identify propositions we limited the participants to a small number of experts.

The second phase was a confirmatory study, which aimed at validating the identified propositions. A total of 30 participants participated in the study while responses from 25 participants (16 researchers and 9 practitioners) were used for the study (see Table 2).

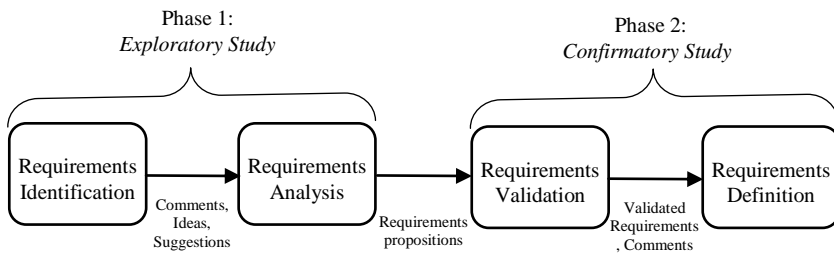


Figure 14: Stakeholders' requirement elicitation process.

The purpose of this study was to validate the propositions by letting a number of participants judge the validity of the propositions. It is therefore, for this reason a larger number of experts were involved in the validation.

Finally, based on the analysis of the data collected from the confirmatory study, requirements for a process model repository are suggested.

Table 2: The study participants

Study type	Method	Total participants	Participant type	No. of participants
Exploratory Study	Interviews	5	Practitioners	2
			Researchers	3
Validation Study	Survey	25	Practitioners	9
			Researchers	16

4.2.1 Exploratory Study

In this section we describe how the exploratory study was conducted. Furthermore, the analysis of interviews is discussed in order to define a set of propositions.

4.2.1.1 Selection of Participants

In order to include both industrial and academic perspectives the study involved both practitioners and researchers. A total of five participants were interviewed. Of the five interviewees, three were researchers and two were practitioners. The researchers are academics who are doing research in business process modelling and enterprise modelling in general. The practitioners are experts who have worked for more than a decade in the business process management industry and have worked on various projects as business analysts or modelling experts.

4.2.1.2 Questionnaire Preparation

The questionnaire was prepared using information gathered from the literature study, personal experiences and our previous research work. The questions were open-ended (see Appendix A.2) so that the interviewees could discuss their experiences, important issues and problems faced when modelling and reusing process models.

4.2.1.3 Conducting the Study

During the interview each respondent was given a copy of the questionnaire to enhance his or her understanding of questions. Each interview lasted for approximately one hour and all the interviews were recorded. Based on the answers of the participants some follow-up questions were asked to pursue interesting issues that came up during the conversation. The interviews were recorded with the participants' knowledge.

4.2.1.4 Analysis

A thematic analysis (Ritchie & Spencer, 2002) was used to code and scrutinize the audio data from the interview to identify key concepts, categories and themes.

The analysis of the interviews indicates that modelling business processes from scratch is difficult and time-consuming. This has also been affirmed in existing studies (Markovic & Pereira, 2008b; Hornung et al, 2009; Rodrigues et al, 2006). Another challenge indicated by the respondents is the difficulties of getting information from the people working in a business process. They agreed that reuse of process models may reduce these challenges, which justifies the need for a process model repository. However, practitioners pointed out that the reuse of process models heavily depends on the quality of the content stored in the repository.

The detailed analysis of the collected data showed that the features and characteristics most desirable for a process model repository to support reuse of process models can be grouped into: *business domain perspective*, *modelling language support*, *business process representation*, *business goals association*, *business process evolution* and *business environment*.

Business domain perspective. Process model reuse implies taking a process model from the repository and using it as a starting point for modelling organization-specific processes. The response from interviews indicates that a lot of effort may be required to redesign a process model that is specific to a particular domain before it can be reused in a different domain. Therefore, a repository that is not limited to domain-specific process models may increase the flexibility of sharing, modifying and therefore reusing process models.

Modelling language support. A number of process modelling languages (e.g. BPMN, EPC and YAWL) are used for modelling business processes. These languages have different constructs, therefore specifications of a business process may vary from one language to another (List & Korherr, 2006). The interviews indicate that users of one modelling language may not understand a process model written in another language. Therefore, a repository that supports different modelling languages may increase the chance of process model reuse.

Business process representation. The description of a business may exist in two forms: textual representation and graphical representation. The responses from the interviews show that while it is easier to store and share processes in textual form, it is much easier for users to comprehend a business process in a graphical representation. Therefore, providing graphical representations of business processes will increase the likelihood of comprehending and reusing them.

Business goals association. A goal is a condition or state of affairs that an actor wants to hold (Edirisuriya, 2009). The purpose of a business process is to achieve one or more goals. The respondents indicate that one of the most important aspects users consider when searching for a business process is whether it can achieve a certain goal or not. Therefore, relating process models with goals in the repository can help users to understand and thereby reuse process models.

Business process evolution. The dynamic and competitive nature of most business environments requires organizations to often change and adapt their

business processes to meet specific business demands (Zhao & Liu, 2007). The respondents indicated that the original and adapted processes should be maintained in the repository for future reuse. In addition to that, the importance of representing a business process with different levels of detail in a repository was discussed by respondents.

Business environment. The response of the interviews indicate that in order to reuse process models it should be possible to easily locate the required process, i.e. through navigating, searching, querying, etc. In addition, it is important to know the *environment* in which a process can or is intended to work. The respondents suggest that this environment consists of the business context in which the business process can be applied, the goals of the process and the actors of the process.

Table 3: The derived propositions

ID	Propositions
P1	A domain expert can understand a process related to its domain written in any process modelling language.
P2	Experts of a process modelling language with common knowledge of a domain can understand any process model in that domain written in a language of his/her expertise.
P3	Reuse of process models can simplify the work of modelling business processes, improve modelling efficiency and reduce the cost of modelling business processes.
P4	A process model written in one process modelling language is difficult to reuse for users of another language.
P5	Domain-independent process models can be reused for modelling specific business processes in an enterprise.
P6	A graphical representation of a business process is easier to understand than a textual representation.
P7	Communication gaps between business experts and IT designers often cause problems in understanding process models.
P8	A process model repository can help reduce the gap between business experts and IT designers.
P9	A process model repository can play an important role in reusing process models.
P10	A repository can support reuse even if only fundamental elements (activities, agents, control flow) of process models are stored, i.e. composite tasks, intermediate events etc. are omitted.
P11	In a repository, it is useful to represent the same process using several process models with different levels of detail.
P12	A repository should maintain multiple versions of all its process models.

4.2.1.5 Propositions

Propositions are the suggestions of the requirements for a process model repository that are derived from the analysis of the interviews. The interviews were carefully transcribed and all the issues discussed by the interviewees were listed, i.e. we were exhaustive in our approach to include

all the suggestions gathered from the experts. Following that, the conflicting statements were omitted and similar statements were grouped and rephrased in such a way that redundancy could be avoided. Table 3 provides the list of propositions.

4.2.2 Confirmatory Study

Since only a small number of participants (five) were involved in the exploratory study, a confirmatory study was conducted to affirm the propositions. In this section, a brief overview of the criteria used for selecting the participants and the study procedure is presented followed by an analysis of the data collected through the empirical study.

4.2.2.1 Questionnaire Preparation

For validating the propositions elicited from the interviews a questionnaire (see Appendix A.3) was prepared. Prior to conducting the study, a pilot study was conducted to check the consistency of the questionnaire. In this case the questionnaire was sent to six participants and responses from four participants (67%) were received. Based on the feedback of the pilot study, the questionnaire was revised, i.e. the language of some questions was improved and one question was removed from the questionnaire. In addition to that, the questionnaire was divided into two sections, one related to the content of the repository and another to the repository itself. The questionnaire had a set of propositions and an evaluation scale from 1 to 5. A score of 1 is for *strongly disagree*, 2 is for *disagree*, 3 is for *not sure*, 4 is for *agree* and 5 is for *strongly agree*. The design of the questionnaire allowed participants to provide additional comments on each question.

4.2.2.2 Conducting the Study

The participants of the study consisted of researchers and practitioners who attended the 2nd IFIP WG 8.1 Working Conference on the Practice of Enterprise Modelling (PoEM'09), Stockholm, Sweden (Zdravkovic, 2009). A total of 37 participants were given the questionnaire, from which 30 participants submitted their completed questionnaire, making the response rate 81%.

During the study, participant information was kept partially anonymous, i.e. personal information (name, email, etc.) about participants was not collected. However, information about area of expertise was collected in order to ensure that only data from process or enterprise modelling experts was collected. Furthermore, role information (researcher/practitioner) was collected in order to distinguish between the responses from researchers and practitioners.

4.2.2.3 Selection of Participants

Only the responses from participants who met the following criteria were included: a) area of expertise is business process and/or enterprise

modelling; b) for researchers the minimum qualification is PhD student. Furthermore, incomplete responses and participants who had not marked their profession (researcher and/or practitioner) were omitted. Therefore, the confirmatory study includes data from 25 participants (9 practitioners and 16 researchers), which is 83% of the response.

4.2.2.4 Analysis

In order to validate the propositions, the data collected from the confirmatory study is analysed and discussed. Furthermore, the analysis of the differences between responses from practitioners and researchers is discussed.

Propositions Analysis

For analysing the propositions, a *frequency distribution analysis* has been used in order to identify the distribution of responses over a scale of 1 to 5 (strongly disagree to strongly agree). Figure 15 shows the frequency distribution of the propositions from P1 to P12. The y-axis represents the propositions whereas the x-axis represents the percentage of the response. The different colours of the bars represent different values (1 to 5) as shown in the graph key. From the graph it is evident that a large number of participants at least agree (either agree or strongly agree) with most of the propositions. Table 4 presents detailed results of the frequency distribution of each proposition and important results are discussed below.

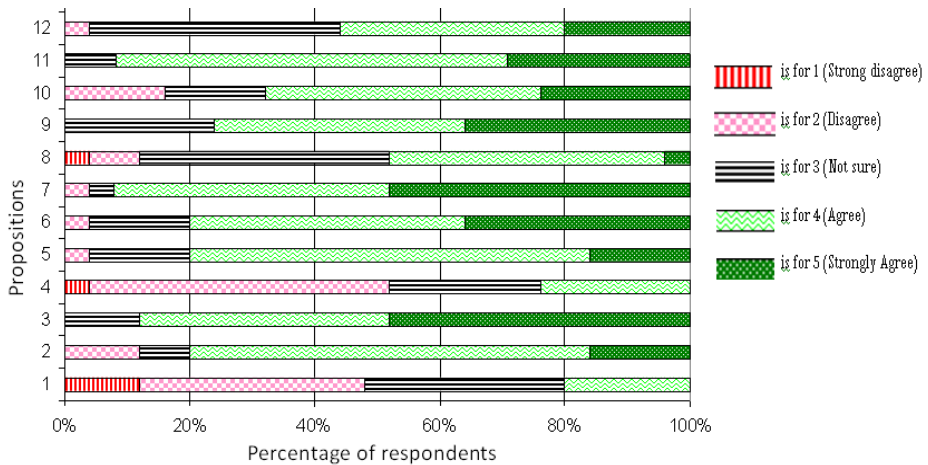


Figure 15: Frequency distributions of propositions (P1–P12).

From Figure 15 and Table 4 it can be seen that a large percentage (80%) of participants at least agrees (either *agree* or *strongly agree*) with P2 (Experts of a process modelling language with common knowledge of a domain can understand any process model in that domain written in a language of his/her expertise). The results indicate that the understanding of a business process

depends upon the knowledge of the process modelling language rather than the knowledge of the domain to which the process belongs.

A proportion of 88% of the participants at least agrees (either *agree* or *strongly agree*) with P3 (reuse of process models can simplify the work of modelling business processes, improve modelling efficiency and reduce the cost of modelling business processes). These results are consistent with the results obtained from the exploratory study (interviews). Furthermore, some participants added that in their experience “*once modelling time of processes is reduced the saved time can be used for optimization of processes*”.

Table 4: Propositions analysis (P1–P12)

Propositions	Strongly Disagree %	Disagree %	Not Sure %	Agree %	Strongly Agree %
<i>P1</i>	12	36	32	20	0
<i>P2</i>	0	12	8	64	16
<i>P3</i>	0	0	12	40	48
<i>P4</i>	4	48	24	24	0
<i>P5</i>	0	4	16	64	16
<i>P6</i>	0	4	16	44	36
<i>P7</i>	0	4	4	44	48
<i>P8</i>	4	8	40	44	4
<i>P9</i>	0	0	24	40	36
<i>P10</i>	0	16	16	44	24
<i>P11</i>	0	0	8	63	29
<i>P12</i>	0	4	40	36	20

A proportion of 52% of the participants at least disagree (either disagree or strongly disagree) with P4 (a process model written in one process modelling language is difficult to reuse for users of another language), whereas 24% are not sure about P4. This means that a process model written in one language is not difficult to reuse for users of another language. Therefore, a process model repository may support reuse for a large percentage of users even if multiple process modelling languages are not supported by the repository.

A large percentage (92%) of participants at least agree (either agree or strongly agree) with P11 (In a repository, it is useful to represent the same process using several process models with different levels of detail), whereas the remaining 8% were not sure. Therefore, the repository should support multiple views of a process (with different levels of detail). The view management functionality for process repositories has also been discussed in existing studies such as the work by Yan et al (2012).

Practitioners and researchers comparison

In this subsection we compare and contrast between the responses from practitioners and researchers to determine any significant differences. This is done by analysing the variance between the responses from practitioners and

researchers by using an ANOVA test with the help of an online calculator available at (Measuring Usability LLC, 2004). The ANOVA for unequal sample size is used for the comparison because the number of practitioners is not equal to the number of researchers. Table 5 presents the results of the analysis, and the main findings are discussed below.

Table 5: Analysis of variance between practitioners and researchers

Propositions	Mean	StDev	P-Value
P1	P = 2.78, R = 2.5	P = 1.09, R = 0.89	0.263
P2	P = 3.89, R = 3.81	P = 0.78, R = 0.91	0.4143
P3	P = 4, R = 4.56	P = 0.87, R = 0.51	0.0511
P4	P = 2.56, R = 2.75	P = 0.73, R = 1	0.2916
P5	P = 4.11, R = 3.81	P = 0.33, R = 0.83	0.1106
P6	P = 4.11, R = 4.13	P = 0.78, R = 0.89	0.4839
P7	P = 4.33, R = 4.38	P = 1, R = 0.62	0.4555
P8	P = 3.22, R = 3.44	P = 0.67, R = 0.96	0.2586
P9	P = 3.78, R = 4.31	P = 0.83, R = 0.70	0.0631
P10	P = 3.78, R = 3.75	P = 1.09, R = 1	0.4751
P11	P = 3.89, R = 4.12	P = 0.33, R = 1.26	0.2442
P12	P = 3.78, R = 3.69	P = 0.67, R = 0.95	0.3921

where P is for practitioners and R is for researchers.

◆ Practitioners
 ◆ Researchers
 x-axis response scale from *strongly disagree* to *strongly agree*
 y-axis percentage of respondents

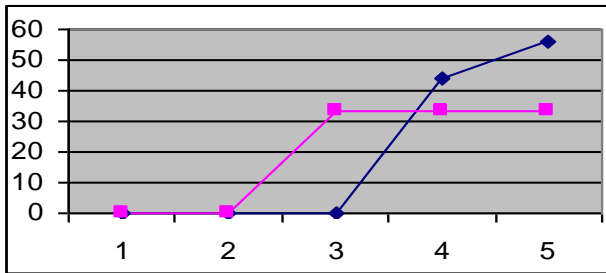


Figure 16: Researchers and practitioners comparison on proposition P3.

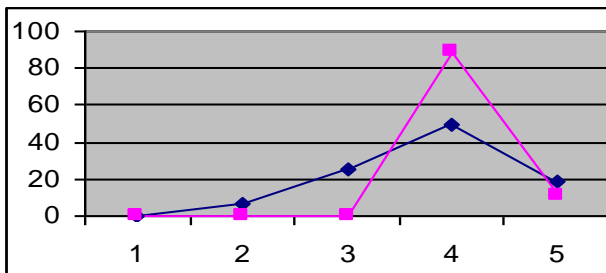


Figure 17: Researchers and practitioners comparison on proposition P5.

Figure 16 shows that there is a significant difference between practitioners and researchers in supporting P3 (reuse of process models can simplify the work of modelling processes, improve efficiency and reduce modelling cost). Furthermore, from Table 5 it can be seen that the deviation in response between practitioners (StDev = 0.87) is less than that of researchers (StDev = 0.51), which implies that researchers are more confident about P3. Therefore, it is likely that the reuse of process models is more accepted by the research community than among practitioners.

Figure 17, shows that there is a significant difference between practitioners and researchers in the acceptability of P5 (domain-independent process models can be reused for modelling specific processes). Furthermore, from Table 5 it can be seen that the deviation in response between practitioners (StDev = 0.33) is significantly less than that of researchers (StDev = 0.83), which implies that practitioners are prepared to adopt domain-independent process models more than researchers.

Table 6 summarizes the requirements identified from the validation of the propositions. The second column refers to the propositions and the third column refers to the key concepts related to the requirements.

Table 6: Requirements (from stakeholders)

RN	Requirement	Propositions	Concepts
R1	The repository should be able to store process models in at least one process modelling language.	P4	Modelling language support
R2	The repository should allow process models to be stored regardless of their domain – storing both domain-specific and generic process models.	P5	Domain independence
R3	Process models in the repository should be represented in both graphical and textual form.	P6	Representation
R4	The repository should store both business and process models.	P7	Business model inclusion
R5	In the repository, a business process should be represented by several process models with different levels of detail.	P11	Granularity
R6	The repository should allow multiple versions of a process model to be maintained.	P12	Versioning
R7	A process model should be annotated with information that can facilitate searching, navigating and interpreting of process models.		Annotation
R8	Process models in the repository should be categorized based on widely accepted classification schemes to facilitate navigation.		Classification scheme

4.3 Requirements Elicitation from Literature Survey

The requirements elicitation was also based on the findings of the literature survey. This method was appropriate, as the new solution would perform the same sorts of tasks as the other repository solutions, except in the way that they support the reuse of process models.

We have applied the three steps of systematic review (Kitchenham & Charters, 2007) in order to obtain requirements for the process model repository from the literature: planning, conduction and requirements establishment. Next, these steps are discussed in more detail:

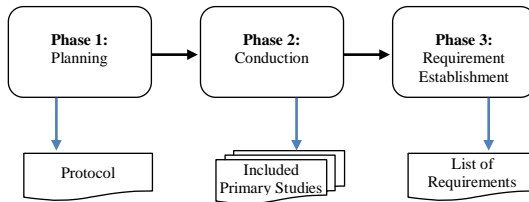


Figure 18: Requirement elicitation – a systematic review process.

4.3.1 Planning

In the first step, a review protocol was prepared. The review protocol comprised the research questions, search strategy, search sources, and criteria for inclusion and exclusion. Suitable research questions are vital to successfully establish the set of requirements. In this research, we were interested in identifying: “requirements for the design of the process model repository for process model reuse”. Therefore our main research question was “which requirements have been considered during the design and development of the process model repositories?”

With regard this research question, we identified the main keywords: “process model repository” and “process model reuse”. Then we found related terms for these keywords: “business process repository” and “business process reuse”. The search source for our systematic review was Google Scholar.

Another task was to define the inclusion criteria (IC) and exclusion criteria (EC). These criteria enables the inclusion of primary studies that are relevant to answer the research questions and exclude irrelevant studies. Thus, the inclusion criteria are:

- i. The title of the primary study explicitly includes or implicitly refers to the area of interest: “business process model repository”.
- ii. The primary study describes or proposes a repository of business processes.
- iii. The primary study describes or discusses reuse of business processes or process models.

Thus, with regard to the review protocol, the conduction of the systematic review followed in the next phase.

Table 7: Selected primary studies

S/N	Authors	Year	Brief Description
S1	Malone et al	2003	Discusses MIT process handbook
S2	Shahzad et al	2009	Requirements for a process model repository
S3	Yan et al	2012	A framework and survey of business process repository
S4	La Rosa et al	2011a	Discusses an advanced process model repository
S5	Jin at al	2013	Discusses a method for effecting querying of large process model repositories
S6	Kunze and Weske	2011	Discusses an indexing approach that saves comparison operations during search
S7	Eid-Sabbagh et al	2012	Discusses requirements, design and implementation for a process library
S8	Qiao et al	2011	Discusses an approach to cluster and retrieve business processes
S9	Jung et al	2009	Discusses methodology of business process clustering based on process similarity
S10	Decker et al	2008	Discusses an online process modelling platform
S11	Fantinato et al	2012	A survey of reuse in the business process management domain
S12	Aldin & de Cesare	2011	A survey of existing literature on the problem of BPM reusability
S13	Holschke et al	2009	Discusses process granularity as a factor in design tasks under reuse

4.3.2 Conduction

In the second step, we conducted a systematic literature search in accordance with the review protocol. We conducted a search of primary studies by observing all studies that matched the search string in the search sources. A total of 67 primary studies were identified. This was followed by selection of the primary studies, through reading of titles and abstracts and applying the inclusion and exclusion criteria. A total of 25 primary studies were considered for full reading as well as application of inclusion and exclusion criteria. Only primary studies that contributed to obtaining requirements of the process model repository were considered. Finally, 13 studies, as shown in Table 7, were considered the most relevant to the requirement elicitation. This table presents the authors, publication year and a brief description of

the primary studies. In the next step, the requirements of the process model repository are extracted based on these primary studies.

Table 8: Requirements (from the literature)

RN	Requirement	Source	Concepts
R9	The repository should provide a mechanism to enable analysis and comparison of similarities between process models for easy retrieval	S4, S5, S6	Similarity search
R10	In the repository the relationship between process models should be defined and maintained to enable users to identify and find process models that are related to a candidate process model	S1, S2	Process Relationship
R11	It should be possible for users to create new process models or edit existing ones online	S10	Process modelling
R12	It should be possible to integrate the repository with external modelling tools, repositories to allow exchange of process models and other process information	S3	Integration
R13	The repository should provide an access control mechanism to enable sharing and exchange of process models	S3, S7	Access control
R14	The repository should accommodate future expansions of the repository content and structure	S2	Extensibility
R15	The repository should provide tools and mechanisms to manage changes to both repository contents structure and functionalities	S3	Configuration management
R16	The repository should provide a mechanism to manage process model granularities	S11, S13	Granularity
R17	Process models in the repository should be clustered to enable easy retrieval	S8, S9	Clustering
R18	The repository should store other information along with process models to enable redesign of process models from higher-level process knowledge	S4, S11, S12	Patterns, fragments
R19	The repository should provide a mechanism to manage process model variants	S3, S11	Versioning

4.3.3 Requirement Establishment

In order to establish the set of requirements of the process model repository for process model reuse, we have again conducted a full reading of each primary study (listed in Table 7). We have extracted the requirements and

written down a set of requirements provided by each study. Subsequently, we have examined each set of requirements and synthesized them in a distinctive set of requirements. Finally, 11 requirements were identified. The requirements are listed in Table 8. The second column refers to the requirements and the third column is related to the sources that contributed to establishing that requirement. For instance, requirement R9 (Similarity) has three sources: primary studies S4, S5 and S6. Furthermore, in order to support the design of the process model repository, the fourth column presents the concept related to each requirement.

4.4 Requirements Specification and Justifications

Primarily based on the analysis of the stakeholder requirements and requirements from literature search we define and justify the following requirements for a process model repository. It is notable that the presented requirements can be extended and adapted based on the primary purpose of the repository.

Requirement 1. Standard process model notation support

The reuse of a process model depends on the ease with which a user can comprehend, transform and adapt it for the specific business needs. The use of standard process modelling notation makes it easier to comprehend, transform and adopt process models for reuse. While there are several standard modelling notations (e.g. BPMN, EPC, YAWL, etc.), the study has shown that users of one modelling notation can understand and reuse process models written in another notation; *therefore a repository should support at least one standard process modelling notation.*

Requirement 2. Domain-independent process model support

The reuse of process models implies taking a process model from the repository and using it as a starting point for modelling organization-specific processes. Process models that are specific to a particular domain may require a great amount of customization before they can be reused in a different domain. A repository that is domain-independent (not restricted to domain-specific process models) will increase the flexibility of sharing, modifying and thus reusing process models. *Therefore the repository should allow process models to be stored regardless of their domain – both domain-specific and generic process models.*

Requirement 3. Process model representation

Graphical notation is considered easier to comprehend than textual notation (Bauer & Johnson-Laird, 1993; Stenning & Oberlander, 1997). In contrast, graphical notation is not as expressive as textual notation since some aspects of business process characteristics cannot be specified completely using only diagrams (Petre, 1995). The combination of graphical notation and textual

notation in specifying a business process may build a synergy between the two notations. *Therefore, process models in the repository should be represented in both graphical and textual form.*

Requirement 4. Business model definition

A business model provides a high-level view of the business activities performed by an enterprise, stating *what* is offered by *whom* to *whom*, whereas a process model focuses on the operational and procedural aspects of these activities – *how* a particular business case is carried out (Bergholtz et al, 2003). Providing both business and process models in the repository may enable users to identify clearly a process model that meets their business requirements. *Therefore, a repository should provide the definition of both the business and process models.*

Requirement 5. Process model granularity

The granularity of process models may vary depending on the need they fulfil (Polyvyanyy et al, 2008; Holschke et al, 2009). Top management require a rather large-grained process description, as they want to have a general overview of the time and resources needed to accomplish the process, whereas developers, users and analysts prefer a fine-grained process model for the details of a business process. *Thus, in the repository a business process should be represented by several process models with different levels of detail.* In addition, it should be possible to view a coarse-grained process model from detailed ones to meet their needs.

Requirement 6. Version management

Business processes may differ across business units within the same organization, across organizations within the same industry or across industries (La Rosa, 2009). Therefore, the same business process may be represented by different process models. Furthermore, the dynamic and competitive nature of most business environments requires organizations continuously to adapt their business processes to new conditions (Bae et al, 2007; Zhao & Liu, 2007). Thus, *the repository should provide multiple versions of process models for the same business process.* In addition, it should be possible to find and view all process model versions of a business process so that they can choose the relevant model.

Requirement 7. Advanced process model search

Providing efficient search instruments is fundamental to enabling users to retrieve relevant process models as quickly as possible. When the size of the repository is very large, it may be difficult for users to find relevant process models. To narrow the area of the search, we need a way to annotate stored process models (Andersson et al, 2005). Such annotation may also enhance users' understanding of process models in order to decide whether to reuse them. *Therefore, process models should be annotated with information that*

would allow users to narrow down the area of search so they can find relevant process models more quickly.

Requirement 8. Repository navigation

Searching requires users to enter relevant keywords or annotation to locate process models (Huang, 2003), whereas navigation (browsing) is a retrieval process that requires users to navigate a repository by following links from one item to another (Huang, 2003). To improve the efficiency of the navigation system, users should be able to navigate the repository content by process categories so that they can find relevant process models more quickly. *Therefore, process models in the repository should be categorized based on well-established schemes to improve navigation efficiency.*

Requirement 9. Analysis and comparison of process model similarities

For a large collection of process models, a retrieved sample of process models may contain very similar components differing only in minor design details. Users of the repository should be able to discriminate very similar process models and select process models that require the least modification effort (Shahzad et al, 2009). *Therefore the retrieval system of the repository should provide a mechanism to analyse and compare among similar process models.*

Requirement 10. Process model relationship

One of the recognized issues in designing the process model repository is the need to define the relationship between business processes in the repository (Malone et al, 1999; Malone et al, 2003; Shahzad et al, 2009). The purpose of defining the relationship between process models in the repository is to allow users to easily find process models that are related to a candidate process in the repository. A typical scenario for using the process relationship service in the repository is in user-initiated navigation tasks where the user retrieves the complete set of relationships for a particular process model, inspects the relationships, selects the targets and retrieves the process model of this target possibly along with the relationships that the target process model participates in.

In the repository the relationship between process models should be defined to enable users to identify and find process models that are related to a candidate process model.

Requirement 11. Process modelling

The ability to reuse existing process models includes activities related to retrieval, adaptation and incorporation. The retrieval step retrieves a list of process models and selects one that meets the business needs. The adaptation step generates the target model, usually by modifying the selected model. The new model is then used to address the specified need. In addition, the incorporation step takes the newly created process model and inserts it back into the repository. As new process models are created and

added to the repository, they in turn can be reused. This requires users to be able to create or modify process models, share them with collaborators or make them available to the public (Decker et al, 2008). *Therefore, the repository should allow process models to be created or modified, without requiring additional software installation for the client side.*

Requirement 12. External repository access

There are several process model repositories. Sharing process models across repositories will increase the probability of reaching a critical mass of process models. Consequently reuse of process models will improve. With the advancement of Web technologies, an increasing number of process model repositories want to share process models over the Web. The most important requirement to achieve this goal is to build a methodology for homogeneous access of process models. *Therefore, it should be possible to access and share process models with external repositories.*

Requirement 13. Security

The main goal of this research is to design an open and language-independent process model repository. Publicly open means users must be able to access and edit stored process models. Security is required to make such interactions possible and to maintain the correctness and consistency of the repository content. *Therefore the repository should provide a security mechanism to control access to the repository and operations on process models to enable sharing and exchange of process models between contributors.*

Requirement 14. Repository capability and information structure extensibility

For a publicly open repository, the functionalities and the information structure are expected to grow. Therefore the repository should be designed to allow and accept any significant extension without major changes to the architecture and rewriting of code.

Requirement 15. Configuration management

Based on the requirements established so far such as extensibility and openness, it is obvious that the proposed repository will keep evolving. In order to ensure that the repository serves the mission of the users who utilize it while accommodating new changes, it is necessary for the repository to provide tools and mechanisms to manage expected changes. *Therefore the repository should provide tools and mechanisms to easily manage changes in the repository.*

Some of the requirements gathered from stakeholders (see Table 6) were similar to or the same as the requirements extracted from the literature (see Table 8). For example requirement R5 (from stakeholders), is the same as R16 (from literature search). They are related to managing process model

granularity. Requirement R6 (stakeholders) is also the same as requirement R19 (literature); it is related to version management. Another requirement is R8 (stakeholders), which is the same as R17 (literature). It is related to providing process classification to enable easy retrieval of process models through navigation and searching.

Requirement R4 (stakeholders) is also very similar to R18 (literature); it is related to storing additional process knowledge that can be reused to design new models. While R4 identifies business models to be included, R18 identifies process patterns. Therefore R4 is modified to include R18 in the requirement specification.

There are also several generic requirements for any repository, for example multi-user environment, consistency management, content management, check-in and checkout, etc. In addition, some of the requirements for the process model repository include those identified (La Rosa et al, 2011a) for an advanced process model repository. They include process model evaluation and design. However, in this study only the specific requirements for a process model repository were addressed.

4.5 Related Work

There have been a few attempts to establish requirements for business process model repositories. One of the early works is that of Shahzad et al (2009), who suggest a set of requirements for a business process model repository to support reuse of process models. This work is part of the process model repository project, which this thesis is also part of. The requirements were elicited from reviewing existing process model repositories. While elicited requirements revealed an understanding of what the repository system should provide, it is highly recommended to establish the requirements from stakeholders as a primary source (Kotonya & Sommerville, 1998). To get a better understanding of the need and the problem, we elicited the requirements from stakeholders (Shahzad et al, 2010). In addition, we have elicited the requirements from the literature using a systematic review approach.

Yan et al (2012) is another comprehensive work that defines a set of requirements for business process repositories. In this work, the requirements were elicited from existing repositories. The main focus of the requirements defined in this work was to guide the design of the reference architecture of a business process repository. Reference architectures are more abstract, and therefore the defined requirements are also of a general nature for building a general-purpose repository. Our requirements are mainly focused on defining the requirements for designing a process model repository for supporting reuse of process models. However, our work benefits from this work; it is one of the main sources of our requirement established through a systematic review of the literature.

4.6 Summary and Discussion

In this chapter we have managed to establish requirements for a business process model repository to support reuse of process models. The requirements have been elicited from the stakeholders as well as from the literature.

Requirements elicitation from stakeholders was done by collecting initial results from an exploratory study (semi-structured interviews) and validating the results through an empirical study (survey). In addition, the differences and commonalities between practitioners and researchers were discussed. Based on the analysis of the empirical and the exploratory study, a set of requirements for process model repositories were suggested.

Requirements elicitation from the existing literature was done through a systematic review approach. The approach consisted of three steps: planning, which aimed at establishing the review protocol; conduction of the review, which aimed at searching for and selecting the primary studies from which requirements were extracted; and requirements establishment, which included the extraction of requirements through analysis and synthesis of the selected studies.

The requirements both from stakeholders and literature were then specified and justified. In this research, our focus was on requirements that must be possessed by a process model repository to increase the probability of process model reuse. Therefore the presented requirements are not necessarily complete and can be extended based on the aim of a repository.

5 A Context-based Process Semantic Annotation Model

5.1 Overview

This chapter presents a context-based process semantic annotation model (CPSAM) for semantically annotating business process models stored in a repository. The purpose of the model is to facilitate the searching of process models, to support navigation of the process model repository and to enhance the understandability of process models. This chapter builds on Paper III as listed in section 1.5 of Chapter 1. The chapter begins by presenting the requirements that must be fulfilled by the annotation model in order to address the retrieval problem. We discuss how the annotation model was developed following a three-step process – identification of potential annotation elements, validation of the elements and construction of the model. We then presents a context-based process semantic annotation model (CPSAM) with the description of each annotation element structured around business process perspectives as discussed in Chapter 2. The chapter also presents an example illustrating how the implemented annotation model can be used to facilitate process model retrieval. The chapter discusses some of the related work. Finally the chapter concludes with some key remarks about the annotation elements that constitute the CPSAM.

5.2 Requirements for the Annotation Model

In accordance with the design science research process (Joahannesson & Perjons, 2012), we have identified a number of requirements for the annotation model. Only by fulfilling these requirements can the annotation model reach its objectives of facilitating search, navigation and understanding of process models in a process repository. The requirements have been drawn from the literature and the overall repository requirements established in Chapter 4. The requirements from the literature are drawn from the principles of indexing and classification (Warner, 2000; Ezran et al, 2002), information retrieval (Ezran et al, 2002) and the conceptual model quality framework (Lindland et al, 1994).

- *High Annotation Consistency.* The annotations produced by different users should be consistent, i.e. two users annotating the same process model should produce the same or similar annotations. This requirement

is related to the pragmatic quality of the model (Lindland et al, 1994). High annotation consistency indicates that there is a common understanding of the annotation model among users, thereby ensuring that annotations based on the annotation model will communicate the same meaning to different people.

- *High Annotation Correctness.* The annotations produced using the annotation model should be correct. This relates to the semantic quality of the model (Lindland et al, 1994). High annotation correctness indicates that the annotation model elements are accurately understood, thereby ensuring they can be used in a meaningful and correct way.
- *Ease of Annotation.* Annotating process models should be perceived as easy. High ease of annotation means that process models can be annotated with minimal difficulty, thereby ensuring that large numbers of process models can be efficiently annotated. This requirement is related to requirement 7 and requirement 8, drawn from the interview and confirmed in the confirmatory study.
- *Ease of Use.* Searching and navigating process models should be perceived as easy by users. High perceived ease of use means that process models in a repository can be searched and navigated with minimal perceived difficulty, thereby ensuring that repository users will appreciate the experience of using the repository. This requirement is related to requirement 7 and requirement 8, drawn from the interview and confirmed in the confirmatory study.
- *High Usability.* Searching and navigating a process model repository should be more effective, efficient and satisfactory when using models annotated by the annotation model. This requirement is related to requirement 7 and requirement 8, drawn from the interview and confirmed in the confirmatory study.
- *Enhanced Understandability.* Annotating models using the annotation model should help users to better understand process models. Better understanding of process models means that users can readily identify process models that meet their business needs. This requirement is related to the principles of information retrieval (Ezran et al, 2002).
- *High Discriminatory Power.* The annotation model should be able to make clear and fine distinctions between process models, thereby ensuring that processes with different qualities are not annotated in the same way. This requirement is related to the principles of indexing and classification (Warner, 2000).
- *Extensibility.* It should be possible to extend and customize the annotation model. High extensibility means that the model can be used for many domains, thereby enhancing its general applicability. This requirement is related to requirement 14, which requires extensibility of the information structure and functionality of the repository.

5.3 Annotation Model Development

The annotation model has been developed in three steps: identification of potential annotation elements, validation of the identified elements and the model construction, as shown in Figure 19 below. This is related to the design and develop artefact activity of the research process shown in Figure 12.

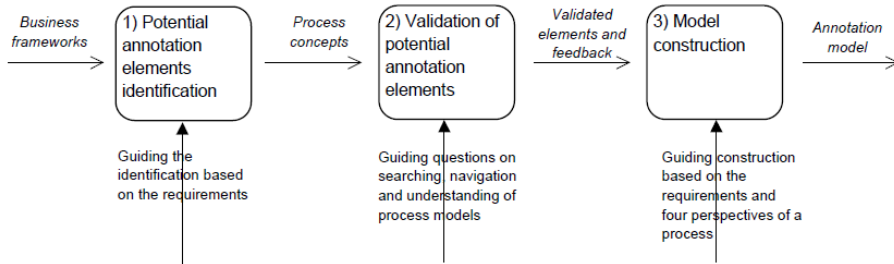


Figure 19: Annotation model development process.

5.3.1 Identification of Potential Annotation Elements

An annotation element is a unit of data that describes one of the properties or characteristics of a business process. In the first step, a set of potential annotation elements was identified by surveying established business frameworks, process classification schemes and business process perspectives. The survey was based on publications in academic as well as trade journals and conferences. In order to identify annotation elements, we searched for relevant journal and conference publications by querying Google Scholar (<http://scholar.google.co.uk>) using the three keyword phrases “process classification scheme”, “business modelling” and “business process modelling”. The results from these searches were narrowed down to publications fulfilling the following criteria:

- i. The title of the publication explicitly includes or implicitly refers to the area of interest, i.e. business modelling, business process modelling, and process classification schemes.
- ii. The publication describes a characteristic or property of business processes, or the publication describes or proposes a well-established business process classification scheme
- iii. The publication has a citation score of at least 5 in Google Scholar

Based on the above criteria, 16 publications were selected, and annotation elements were identified. Studied process classification schemes were included as annotation elements in this step. The identified elements include:

- a) *Process description* (Malone et al, 2003; Guerin, 2005)
- b) *Business context* as defined by (UN/CEFACT, 2001)

- c) *Business goal* from business process perspectives (Kueng & Kawalek, 1997; Lin & Sjølvberg, 2007)
- d) *Domain-specific classification scheme* based on Supply Chain Council (2003) and TM Forum (2009)
- e) *Generic classification scheme* based on Porter's Value Chain (Porter, 2008), the Open-EDI framework (UN/CEFACT, 2003) and the process classification framework (American Productivity and Quality Center, 2001)
- f) *Process property* (Armistead et al, 1995)
- g) *Resource* from REA (Geerts & McCarthy, 2000; Dunn et al, 2005) and the process design framework (Curtis et al, 1992)
- h) *Actor* from REA (Geerts & McCarthy, 2000; Dunn et al, 2005) and the process design framework (Curtis et al, 1992)
- i) *Process relationship* (Malone et al, 2003; Guerin, 2005).

5.3.2 Validation of Potential Annotation Elements

In this second step, the identified annotation elements were validated through a confirmatory study. The activities of the confirmatory study included: preparation of the questionnaire, selection of participants, conducting the study and analysing the results of the study.

Questionnaire preparation. For validating the identified potential annotation elements a questionnaire (see Appendix A.4) was prepared. The questionnaire consisted of a list (from a to i) of the annotation elements as shown above and an evaluation scale from 1 to 5. A score of 1 is for *strongly disagree*, 2 is for *disagree*, 3 is for *not sure*, 4 is for *agree* and 5 is for *strongly agree*. The design of the questionnaire allowed participants to provide additional comments or possible annotation elements on each question.

Participant selection. The participants included 25 volunteer researchers and practitioners who participated in the *2nd Working Conference on the Practice of Enterprise Modelling (PoEM'09)*. Initially 37 participants were involved in the study, however only 30 responded to the questionnaire. From the 30 responses, only 25 complete responses were considered for the validation.

Conduction of the study. The participants were asked to assess whether annotating business processes with the identified elements would facilitate the searching, navigating and interpreting of process models in a repository. During the study, participants were given enough time to provide their comments regarding the annotation elements, and propose additional annotation elements.

Analysis and results. To validate the annotation elements, the data collected were analysed and discussed. The results of the study are displayed in Figure 20. The y-axis represents the annotation elements from a to i, while the x-axis represents the percentage of the response. Different colours

of the bars represent different values (1 to 5) as shown in the graph key. The graph reveals that a large number of participants agree (either agree or strongly agree) with most of the elements.

From Figure 20 it can be seen that most of the participants were not sure of whether generic classification schemes (e) would facilitate searching, navigation and interpretation of process models. *Open EDI phases* and *Porter Value Chain* are the examples of generic classification schemes. A detailed analysis revealed that there is a significant difference between researchers and practitioners in the ability of generic classification schemes to facilitate the search, navigate and interpret process models. Researchers support the use of generic classification schemes more strongly than practitioners. However, in the literature the use of classification schemes for categorizing business processes is advocated (Shahzad et al, 2009; UN/CEFACT, 2005; Gao & Krogstie, 2009) as a means to facilitate search and navigation. Therefore the generic classification schemes were considered for the construction of the annotation model in the next step.

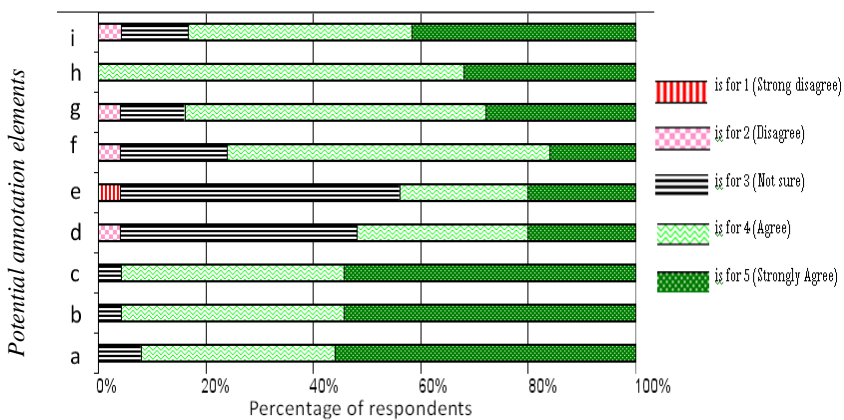


Figure 20: Validation of Concepts.

5.3.3 Model Construction

In this step, based on the feedback from the second step, the semantics of annotation elements and their relationships were defined to form the annotation model.

Business context, goal, resource, actors and process relationship were directly included as elements in the proposed context-based process semantic annotation model (CPSAM). From the generic classification scheme the following elements were defined: *process area* (based on Porter's Value Chain (Porter, 2008)) and *process phase* (based on Open-EDI (UN/CEFACT, 2003)). In addition, we have introduced two further elements that do not have direct relationships with those in step 2. Instead, these

elements were derived from the participants' feedback and their references to different studies. These elements are *process type* (based on the REA ontology (Geerts & McCarthy, 2000; Dunn et al, 2005) and *process level* (based on organizational theory (Anthony et al, 1995)).

5.4 CPSAM – Context-Based Process Semantic Annotation Model

In this section, a context-based process semantic annotation model (CPSAM) is described (Elias & Johannesson, 2013). CPSAM consists of the following annotation elements: “*process type*”, “*process area*”, “*resource*”, “*actor*”, “*organizational level*”, “*process phase*”, “*process relationship*”, “*business context*” and “*business goal*”, as shown in Figure 21. The annotation model extends existing process classification schemes (Porter, 2008; UN/CEFACT, 2003) by incorporating elements from well-established frameworks in accounting (Geerts & McCarthy, 2000; Dunn et al, 2005), organizational theory (Anthony et al, 1995; Anthony, 1995) and enterprise modelling (Huat Lim et al, 1997; Fox et al, 1996).

The annotation elements are structured around five perspectives of business processes discussed in Chapter 2. The first four of these perspectives are those introduced by Curtis et al (1992) to capture the many facets that exist in business process environments, including humans, information and technology. They are functional, behavioural, organizational and informational:

- The *functional perspective*. This perspective focuses on which activities are performed in a process and how these transform resources from input to output. The functional perspective addresses the question “What is to be done?”
- The *behavioural perspective*. This perspective focuses on the timing of activities, i.e. when activities are to be performed, as well as the ordering and control flow of activities. The behavioural perspective addresses the question “When and how will it be done?”
- The *organizational perspective*. This perspective focuses on where and by whom in an organization activities are to be carried out, in particular which agents are responsible for certain activities. The organizational perspective addresses the question “Where will it be done and who is going to do it?”
- The *informational perspective*. This perspective focuses on information artefacts produced, used, modified or exchanged in a process. These artefacts may take the form of documents, electronic messages, database records, etc.

The fifth perspective is the *business process context perspective*, introduced by List and Korherr (2006), which captures other aspects (i.e. business

goals) of a business process. In the following subsections, these five process perspectives are used to group and describe the annotation elements of CPSAM.

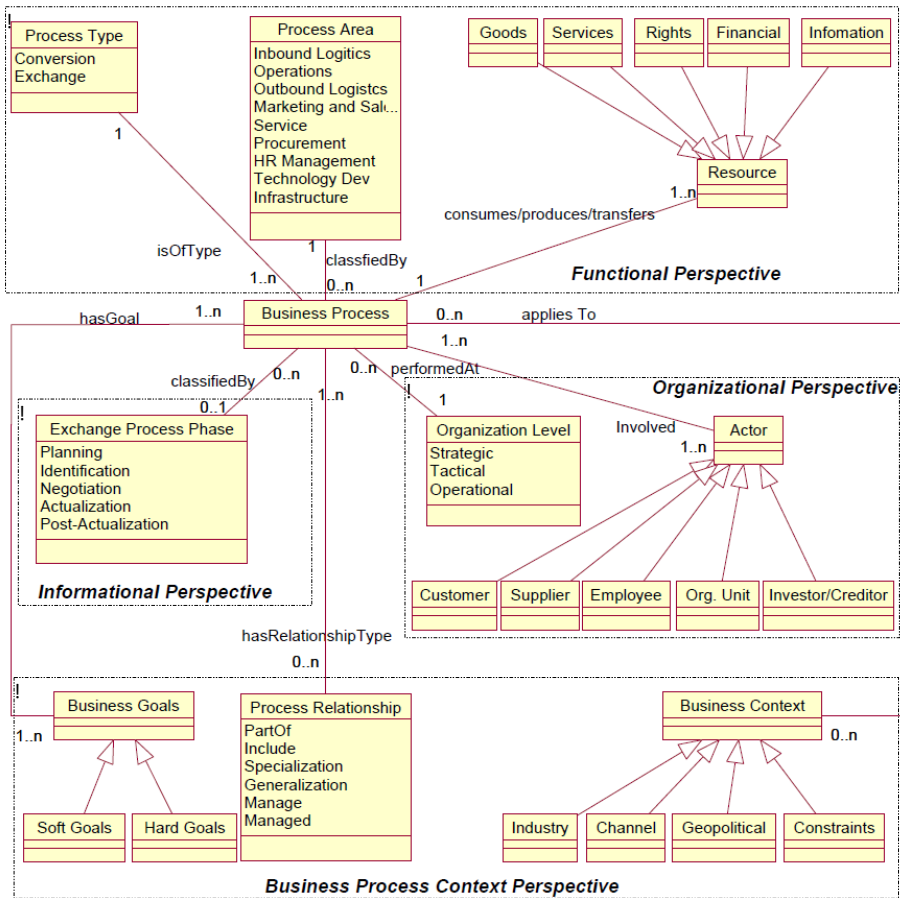


Figure 21: A context-based process semantic annotation model (CPSAM).

5.4.1 Functional Perspective

The functional perspective focuses on activities, resources and the ways in which resources are affected. On a top level, a distinction can be made between processes that aim at exchanging resources between actors and processes that aim at transforming resources:

5.4.1.1 Process Type

The process type element is based on the REA (Resource-Event-Agent) ontology (Geerts & McCarthy, 2000; Dunn et al, 2005). The core concepts in the REA ontology are Resource, Event and Agent. One intuition behind

the ontology is that every business transaction can be described as an event where two agents exchange resources. The process type element classifies business processes according to whether they aim to exchange or transform resources. The possible values of the process type element are:

- *Exchange*: a process in which an enterprise receives resources from another actor and provides resources to that actor in return. The aim of an exchange process is to acquire and pay for resources needed by an organization as well as to sell and deliver goods and services to customers and collect payment. An example is “Sales Process”, which involves exchange of products or services for cash.
- *Conversion*: a process in which an organization uses or consumes resources in order to produce new or modify existing resources. The aim of a conversion process is to convert acquired resources into goods and services for customers. The inputs are transformed into finished goods and services by such a process. An example is “Manufacturing Process”.

While the process type element offers a coarse classification of the function of a business process, the process area element offers a more fine-grained categorization by addressing the value chain of an organization:

5.4.1.2 Process Area

The process area element is based on Porter’s Value Chain (Porter, 2008; Porter, 1985). In order to better understand and distinguish the activities through which an organization creates value, business processes are grouped into process areas. The process area element classifies business processes by their function in the value chain or core competence. A process area can either be primary or supporting. Primary processes consist of activities that create customer value and provide organization distinctiveness in the marketplace, while support processes facilitate accomplishing the primary activities.

The primary process areas include:

- *Inbound logistics*: a process area that includes activities needed for receiving, storing, inventory control, and transportation scheduling.
- *Operations*: a process area that includes activities needed for value creation that transforms inputs into outputs. These include machining, packaging, assembly, equipment maintenance, testing, and other value-creating activities that transform inputs into final products.
- *Outbound logistics*: a process area that includes activities required for getting the finished product to the customers: warehousing, order fulfilment, transportation, and distribution management.
- *Marketing and sales*: a process area that includes activities associated with getting buyers to purchase products including channel selection, advertising, promotion, selling, pricing, and retail management.

- Service and maintenance: a process area that includes activities that maintain and enhance the value of products, including customer support, repair services, installation, training, spare parts management, and upgrading.

The supporting process areas include:

- Procurement: a process area that includes activities needed for acquiring raw materials, services, spare parts, buildings, machines, etc. These include information gathering on resource needs and supplier offerings, supplier contacts, background reviews on the quality of supplier offerings, negotiation, fulfilment, and supplier performance evaluation.
- Technology development: a process area that includes activities that support the value chain activities by developing new technology and procedures, such as research and development, process automation, design, and redesign.
- Human resource management: a process area that includes activities associated with recruiting, development (education), retention, and compensation of employees and managers.
- Firm infrastructure: a process area that includes activities related to general management, planning management, legal issues, finance, accounting, public affairs, quality management, etc.

The elements Process Type and Process Area focus on activities and the ways in which these transform resources. However, from the functional perspective, there is also a need to address the kinds of resources that are exchanged or converted in value chains. For this purpose, the element Resource is introduced:

5.4.1.3 Resource

A resource is an object that is valuable for some actors (Dunn et al, 2005). This element is based on the REA ontology (Geerts & McCarthy, 2000) and the process design framework (Curtis et al, 1992). Resources can be classified in many different ways, therefore CPSAM does not propose one single resource classification. Instead, CPSAM suggests a number of high-level resource categories that can be specialized and complemented with additional categories depending on the domain under consideration. The following resource categories are based on (Dunn et al, 2005):

- Goods: physical or tangible resources, e.g. cars, refrigerators and cellphones.
- Services: abstract resources offered by actors to increase the value of some other resources, e.g. haircuts and eye treatments.
- Rights: privileges, claims or powers of an actor, usually of a legal nature, e.g. ownership rights, usage rights and copyrights.
- Financial: funds or money, e.g. in the form of cash, cheque, voucher or credit card.

- Information: data in a certain context, e.g. blueprints, referrals and customer data.

5.4.2 Organizational Perspective

The organizational perspective includes the actors and roles performing the activities of business processes.

5.4.2.1 Actor

An actor is an entity such as a person or an organizational unit involved in the realization of a business process (Dunn et al, 2005). This element is inspired by REA (Geerts & McCarthy, 2000) and the process design framework (Curtis et al, 1992). Similarly to resources, there is a large number of different kinds of actors and roles. Therefore, CPSAM does not propose one single classification of actors and roles but rather offers a small number of high-level actor categories that can be expanded and specialized, as needed, depending on the domain under consideration. The following kinds of actors are suggested based on (Dunn et al, 2005):

- Customer: an individual, company or organization that buys goods or services.
- Supplier: an individual, company or organization that provides goods or services to a recognizable customer.
- Employee: an individual who provides labour to an organization or another person.
- Investor or creditor: a person, company or entity that puts money or assets into an investment to yield returns.
- Organizational unit: a subdivision or department in an organization that is involved in a business process.

In addition to actors, the organizational perspective also includes the organizational level on which processes exist, which results in the element Organizational Level:

5.4.2.2 Organizational Level

Levels are introduced in organizations in order to allow efficient management and coordination. Most organizations can be seen as operating at three levels: strategic, tactical and operational (Anthony et al, 1995; Anthony, 1995). The Organizational level element describes the level in the organization at which a business process is performed.

- Operational: A business process is said to be at the operational level if it includes activities that are performed on a day-to-day basis. The aim of such a process is to modify and exchange economic resources.
- Tactical: A business process is said to be at the tactical level if it includes activities that are performed in a short-term plan. The aim of such a process is to manage operational-level processes.

- Strategic: A business process is said to be at the strategic level if it includes activities that are performed in a long-term plan. The aim of such a process is to define process types at the operational and tactical levels as well as the resource types to be used and produced.

5.4.3 Informational Perspective

Any business process will produce various information artefacts, but exchange processes in particular produce a large amount of complex business documents and other information. The reason is that exchange processes establish and depend on social and legal relationships between actors that need to be carefully documented. One way to identify and structure this information is by studying the phases of exchange processes:

5.4.3.1 Exchange Process Phase

The activities needed for exchanging resources will extend over several phases from initial planning to follow-up activities after the actual exchange has been completed. In order to classify these activities, the open-EDI framework (UN/CEFACT, 2003) has suggested five process phases, which constitute the basis for the element Exchange Process Phase. Each process phase results in specific information artefacts representing relationships between actors that are used to coordinate their work.

- Planning: The planning phase includes all activities needed to decide what actions to take for acquiring or selling goods and services. Here actors are concerned with the question of what goods or services to acquire or sell. The phase will result in planning documents specifying expected purchases and sales.
- Identification: The identification phase includes all activities needed to identify, select and establish linkages with partners that can be involved in business collaborations. The phase will result in the documentation of potential collaboration partners.
- Negotiation: The negotiation phase includes all activities needed to establish a contract and related commitments for the exchange of goods and services. The phase, when successful, will result in contract documents.
- Actualization: The actualization phase includes all activities needed to prepare and perform the resource exchanges stipulated in the contract established in the negotiation phase. The phase will result in acknowledgements of receipts and other documents testifying to the fulfilment of agreed commitments.
- Post-actualization: The post-actualization phase includes the follow-up activities of resource exchanges performed in the actualization phase, e.g. warranty coverage, complaint handling and after-sales service.

5.4.4 Business Process Context Perspective

In addition to the elements introduced above, there are three more elements, one for expressing relationships among business processes, one for representing their business context, and one for representing process goals.

5.4.4.1 Process Relationship

Business processes interact with each other throughout an enterprise, for example, via outputs from one process forming the inputs for another process (Elias & Bider, 2014). Each process is, therefore, part of a larger whole, and the enterprise can be seen as complex networks of related processes (Elias & Bider, 2014). Understanding how these business processes relates to each other is very fundamental to supporting reuse. The process relationship element describes how business processes are related. The following process relationships have been identified (Malone et al, 2003; Guerin, 2005):

- A partof-includes relationship exists if one process is composed of one or more processes (called subprocesses). The subprocess has the partof role and the parent process has the includes role. An example is “Manage order approval”, which includes “Handle rejected order”, and the latter is a part of the former.
- A generalization-specialization relationship exists if one process (called a specialization) is a kind of another process (called a generalization). An example is “Manage returns”, which is a generalization of “Manage returns with prior approval” and “Manage returns without prior approval”.
- A manage/managed relationship exists if one process plans, controls, monitors, evaluates and/or designs another process.

5.4.4.2 Business Context

In order to reuse a process model, users need to understand the business environment in which it is intended to work. The business environment can best be described by the concept of business context (Hofreiter & Huemer, 2006). A business context defines the circumstances in which a business process may be used (UN/CEFACT, 2001). This element enables users to identify business processes that may only apply to a specific business environment. The context in which a business process takes place can be specified by a set of categories and their associated values (UN/CEFACT, 2001). In CPSAM we define the following contextual categories:

- Industry: the industry in which the business process takes place.
- Communication channel: the channel through which involved actors communicate.
- Geopolitical: provides aspects related to region, nationality or geographically based cultural factors.
- Official constraints: aspects of the business situation that result from legal or regulatory requirements.

5.4.4.3 Business Goal

The purpose of a business process is the achievement of one or more goals. A goal is statement about a condition or state of affairs that an actor would like to achieve (Kueng & Kawalek, 1997). The main concern of business analysts, when modelling business processes, is the achievement of new business goals by the new design. Therefore in order to reuse a process model, the analyst needs to understand the business goals that are achieved by the existing process models. This element describes the business goals that a process model is intended to achieve. According to (Lin & Sølvsberg, 2007), there are two types of goals associated with a business process:

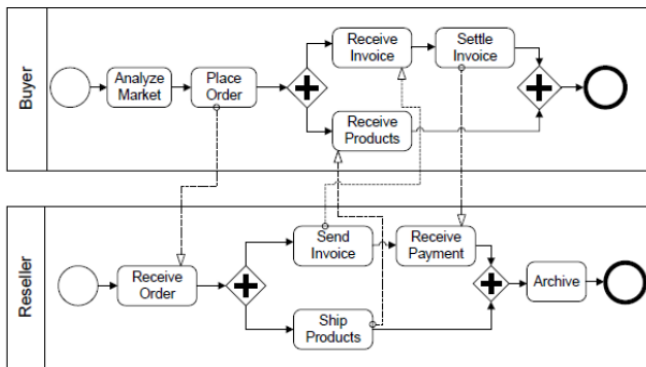
- Soft goals: These are strategic goals that an organization is striving to achieve. For example, a soft goal for a “procurement process” could be “minimize procurement costs”.
- Hard goals: These are operational goals that define the state to be reached by a process (e.g. “complete an order”).

5.5 Demonstration of the CPSAM Application

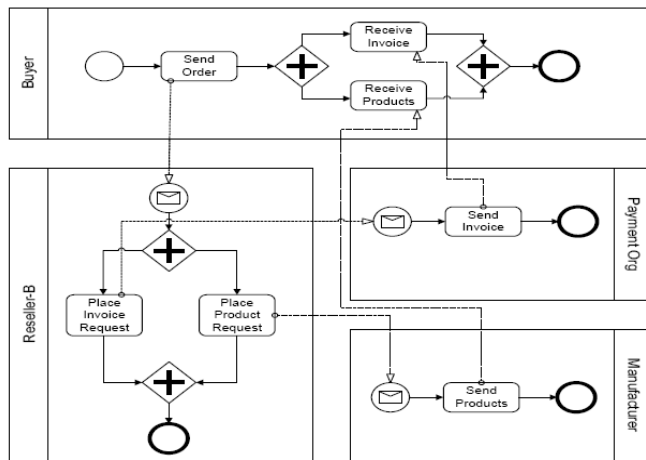
In accordance with the design science research process (Joahannesson & Perjons, 2012), the use of the artefact must be shown in one or a number of cases to prove its feasibility. In this section we demonstrate the use of a semantic annotation tool (see Chapter 6 for the details of the tool), that implement the annotation model, for annotating and retrieving process models. This is related to demonstrate artefact activity of the research process shown in Figure 12. For the demonstration an order-to-cash business process is used as a running example. Order-to-cash is a process where goods are ordered, delivered and received, as well as invoiced and paid for.

All order-to-cash processes include activities related to invoicing, delivery and payment, but they may have several differences. For example, an order-to-cash process that involves the delivery of services, such as training services, is different from the one for the delivery of goods, such as computers. The use of CPSAM for classifying and describing these processes captures their similarities and differences, thus enabling a repository user who is searching for one of the two processes to find a relevant process model.

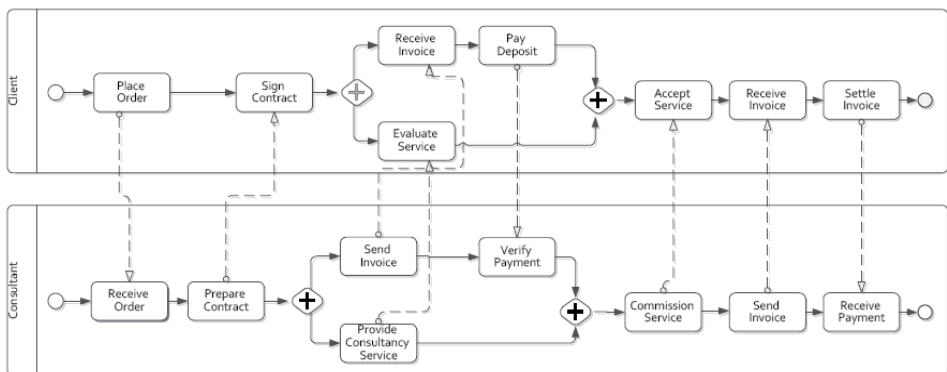
An example scenario begins with a business analyst who uses the annotation tool to annotate an order-to-cash process model, for the delivery of goods, and stores it in the repository. The scenario is followed by another user who performs a search in the repository to find an order-to-cash process model for delivery of service.



(a)



(b)



(c)

Figure 22: Order-to-cash business processes.

Task 1: Annotation of a Process Model

In this task we demonstrate how a user uses the CPSAM to annotate a process model. Suppose the analyst has already designed the order-to-cash process model shown in Figure 22 (a). Before storing the process model, the analyst is required to annotate it:

The analyst is the owner of the process, so he provides the *process name* and *process description* of a process model. The *version number* is assigned automatically.

- The main activities of this process regard getting buyers to purchase products, i.e. selling, therefore the “Process Area” is “Marketing and Sales”.
- This process includes day-to-day activities, therefore “Organization Level” is “operational”.
- The process involves exchange of resources, therefore “Process Type” is “Exchange”.
- The process includes activities for preparing and performing the exchange, therefore “Exchange Process Phase” is “Actualization”.
- The actor involved in the exchange identified as “Principal Actor” is “Supplier” whose role is a reseller, and “Other Actor” is “Customer” whose role is a buyer.
- The reseller receives payment (i.e. cash or cheque) and ships products (goods) to the buyer, therefore the resource being exchanged identified as “Resource Received” is “Financial” and “Resource Provided” is “Goods”.
- The goal of this process is to “Increase customer satisfaction and retention”.
- This is a generic order-to-cash process, not restricted to any domain, therefore it is not annotated with business context information.

The analyst will produce the annotation as shown in Figure 23. This figure is a screen dump from a prototype (described in Chapter 6) built for evaluating the annotation model.

Task 2: Searching a Process Model

In this task we demonstrate how a repository user uses the CPSAM-based annotation to search for an order-to-cash process model (for delivery of service) in the repository. Individuals who model processes, as well as those who formulate the search queries, may use different a vocabulary to express the same concepts. Therefore it is difficult for the user to find the relevant process model by only using keywords.

Using an annotation-based search, the user looking for an order-to-cash process model for consultancy service delivery can limit the search to annotation elements. Starting with *process area*, *organization level*, *process*



Figure 23: Annotation of process model shown in Figure 22 (a).

type, process phase, principal actor and *other actor* to form query 1 shown in Figure 25, the result of query 1 consisted of three process models (shown in Figure 22). Thereafter, a user can proceed with a stepwise strategy, narrowing the query by using more annotation elements. Since the user is looking for an order-to-cash process model for service delivery, this means the *resource provided* in exchange is the “Service” and the *resource received* is “Financial”. This way we decrease the search space. The process model from Figure 22 (c) is retrieved as the relevant process following the execution of search query 2, as shown in Figure 25. Also the figures are screen dumps from a prototype (described in Chapter 6) built for evaluating the annotation model.

The screenshot shows the Activiti Explorer web application interface. The browser address bar displays the URL: `http://localhost:8084/activiti-explorer/ui/#psearch/`. The application header includes navigation tabs for **Tasks**, **Processes**, **Reports**, and **Manage**. The main content area is titled "Search Business Processes" and is divided into two sections: a search filter section on the left and a search results section on the right.

Search Filter Section:

- Process Area:** Marketing and Sales
- Process Type:** Exchange
- Ex. Process Phase:** Actualization
- Organization Level:** Operational
- Principal Actor:** Supplier
- Other Actor:** Customer
- Resource Received:** (empty)
- Resource Provided:** (empty)
- Goal Perspective:** (empty)
- Business Domain:** (empty)

Buttons for "Search" and "Reset" are located at the bottom of the filter section.

Search Results Section:

The search results are displayed in a table with the following data:

PROCESS NAME	PROCESS DESCRIPTION	PROCESS VERSION	PROCESS MODEL
Order.to.cash	An order to cash include activities for ordering, delivering, invoicing and payment	1	View.Process.Model
Order.to.cash	The process include activities for ordering, delivering, invoicing, and payment	3	View.Process.Model
Order.to.cash	This order to cash process include activities for service request, service delivery, invoicing, and payment	2	View.Process.Model

The bottom of the application shows a status bar with the text "© Activiti.org. All rights reserved." and a system tray displaying the time "9:49 AM" and date "1/24/2015".

Figure 24: Search result of query 1.

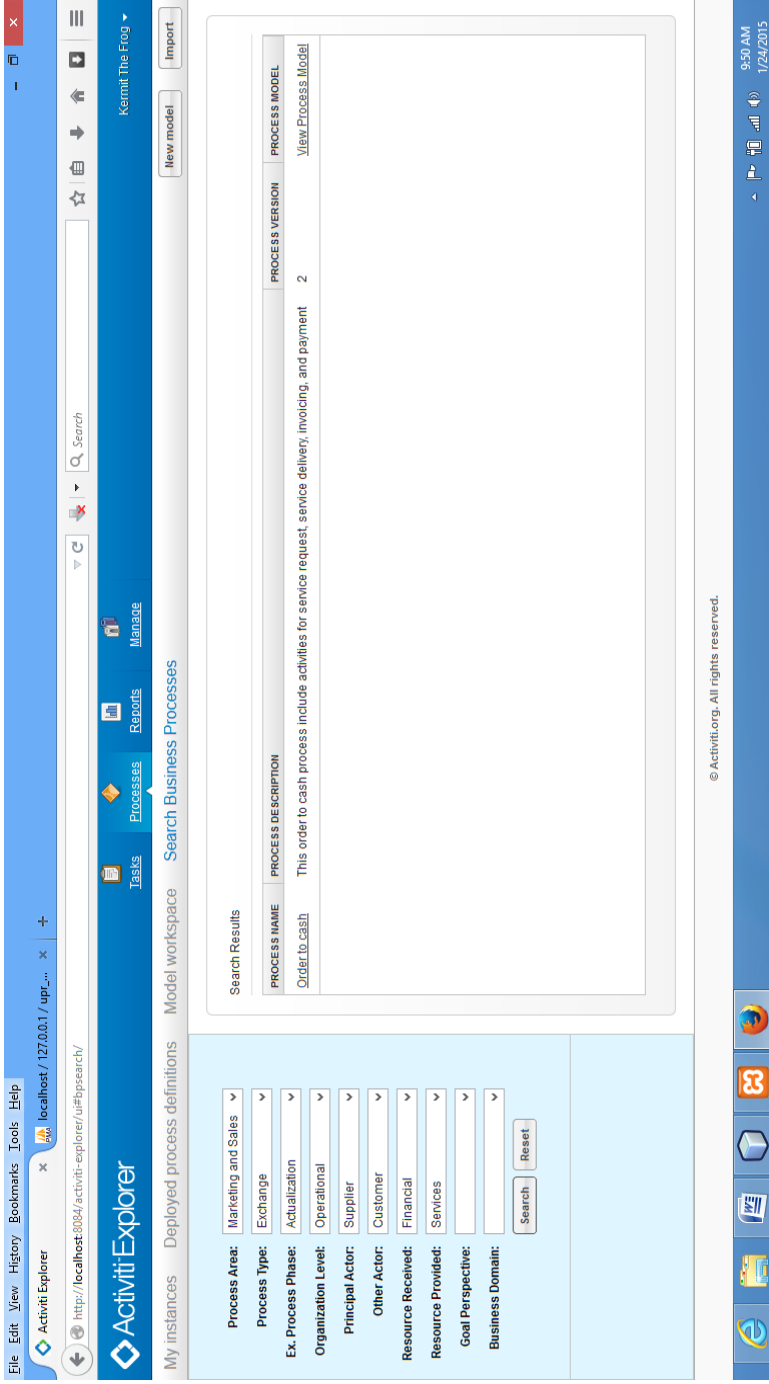


Figure 2.5: Search result of query 2.

5.6 Related Work

The effort to address semantic annotation of process models in the research literature is not new. Lin (2008) is one of the more comprehensive approaches to address the issue of the semantic annotation of process models. In this work she suggests a semantic annotation framework to manage the semantic heterogeneity of process models. Her framework incorporates three perspectives: the basic description of process models (profile annotation), process modelling languages (meta-model annotation) and process models (model annotation). The framework results in a process semantic annotation model that provides a common semantic annotation schema for annotating semi-structured IS solutions. In addition, the semantic annotation model is extended by incorporating goal ontology to specify the organizational objectives.

SUPER project (Wetzstein et al, 2007; SUPER, 2007) is another approach to address the issue of semantic annotation of process models. SUPER aimed to bridge the gap between business and IT, and enables the semi-automation of the BPM life cycle. In this work three main groups of ontologies are proposed for semantically annotating a business process: process, organization-related and domain-specific ontologies (Filipowska et al, 2009). Process ontologies provide a description of the structure of a process such as control flow; organization-related ontologies provide a description of artefacts that are consumed by or involved in the process such as a description of actors, resources, systems, etc.; and domain ontologies provide information specific to an organization from a given domain (Filipowska et al, 2009).

Another approach to address the issue of the semantic annotation of process models is that by Born et al (2007), who proposes an approach for integrating semantics in modelling tools to support the graphical modelling of business processes with information derived from domain ontologies. The proposed semantic information includes objects relevant for each activity as well as the states of these objects and preconditions and post-conditions for the activities within a process model. In order to support users in modelling semantically annotated process models, matchmaking functionalities are defined by matching elements of the graphical business process model with elements of domain ontologies.

Our work differs from those described above in the respect that we focus on using annotated process models for the purpose of repository search and navigation. In contrast, other approaches have a wider scope and also intend to support the design of process models. As a consequence, these approaches require that process models be annotated with specific domain ontologies, typically tailored for the application and domain under consideration. For example, in Lin (2008) individual process fragments and constructs are

annotated with domain-specific annotation elements. Such an approach requires a great deal of effort from both the annotator and the user, who have to learn and apply a specific ontology. CPSAM offers a lightweight approach, where well-known business and process frameworks are used as the basis for the annotation, which provides ease of annotation as well as ease of use. Nevertheless, the extensibility of CPSAM enables domain-specific notions and elements to be included if desired.

5.7 Summary and Discussion

In this chapter we have proposed a process semantic annotation model, which can be used to semantically annotate process models in the repository for an effective and efficient retrieval system. This work builds on the requirements we introduced in Chapter 4 for process model repositories to support process model reuse, and the review and analysis of existing repositories to identify the challenges that limit their use in practice. One of the main challenges faced is that repositories often lack effective instruments for searching and navigating their content. To address this challenge, a context-based process semantic annotation model (CPSAM) has been developed to facilitate searching and navigating models in a process model repository. In addition, CPSAM intends to enhance the user's understanding of process models.

The semantic annotation model is made up of a several annotation elements, where each element describes one aspect of the business process. The annotation model defines a collection of terms or values for each annotation elements. The model uses the predefined values or terms for each annotation elements to describe and represent a process model in the repository in the form of process annotations. The creation of the process annotations is basically the selection of a several of terms or values from each annotation element.

An exemplar application of the annotation model (CPSAM) has demonstrated how the model facilitates improved process model retrieval. A prototype of the semantic annotation tool that implements the annotation model will be discussed in Chapter 6. An evaluation of the model to test its performance and user perception will be discussed in Chapter 7.

6 A Prototype of the Semantic Annotation Tool

6.1 Overview

This chapter is devoted to the prototype implementation of the semantic annotation model concepts discussed in the previous chapter. The purpose of the prototype is twofold: first, to serve as a proof of concept for the implementation and application of the proposed semantic annotation model as the main artefact; second, to enable the evaluation of the semantic annotation model to test whether the model improves the retrieval of process models. The chapter is organized in three parts. It begins with a short overview of the functionalities provided by the prototype system, which is the semantic annotation tool. Then, the chapter presents the prototype system by describing the architecture and the components of the systems. Screenshots showing the core functionalities of the prototype are also provided. A summary and discussion are presented to conclude this chapter.

6.2 Functionalities of the System

The main focus of this prototype is to implement the semantic annotation tool. The semantic annotation tool is based on the semantic annotation model discussed in Chapter 5. The following are the functionalities (use cases) provided by the prototype system:

- The system provides a Web-based process-modelling tool to enable users to model business processes using the browser.
- The system enables users to annotate process models before they are saved in the repository. The annotation is based on the semantic annotation model (CPSAM) discussed in Chapter 5.
- The system provides advanced search capability that is based on process annotation.
- The system provides a navigation mechanism that is based on the annotation model (CPSAM). It also includes an alphabetic-based navigation mechanism.
- The system allows users to import process models, which are either in XML format or an image file. The imported models are annotated before

they are stored in the repository. In addition, the repository also allows the export of process models.

- The system allows users with administrative roles to define, modify and extend the annotation model.

6.3 Activiti

To provide the above functionalities we have implemented the semantic annotation tool within the Activiti framework (Team, 2010; Rademakers, 2012). Activiti is a Web-based BPMS for designing, modelling, executing, optimizing and monitoring business processes; it provides basic functionalities provided by process model repositories. Such functionalities include process modelling, process model storage, sharing, versioning and retrieval. In addition, Activiti is an open source and therefore it can be adapted and customized for specific purposes. It is due to these reasons that Activiti was chosen as a framework for implementing the semantic annotation tool to improve retrieval of process models. It is made up of several different components: Activiti Engine, Activiti Modeler, Activiti Designer, Activiti Explorer and Activiti REST.

- Activiti Engine is the core component of the Activiti framework that performs the process engine functions. It provides the core capabilities to execute Business Process Model and Notation (BPMN) 2.0 processes and create new workflow tasks.
- Activiti Modeler is a Web-based process modelling tool. With Activiti Modeler business analysts are capable of creating BPMN 2.0-compliant business process models. Therefore business processes can easily be shared – no client software is needed before you can start modelling. Activiti Designer is an Eclipse-based plug-in, which enables a developer to enhance the modelled business process into a BPMN 2.0 process that can be executed in the Activiti process engine.
- Activiti Explorer is a Web application that can be used for a wide range of functions in conjunction with the Activiti Engine. It enables a user to get an overview of deployed processes and interact with the deployed business processes.
- The Activiti REST component provides a Web application that starts the Activiti process engine when the Web application is started.

6.4 The Prototype System

Figure 26 presents the architecture of the prototype system. To implement the semantic annotation tool we have used and extended different components of the Activiti stack. We have used *activiti modeler* to provide

the Web-based process modelling functionality. Also the modeller is extended to implement the *annotation service* for annotating process models with semantic information based on the CPSAM. The *retrieval component* of the tool is implemented by extending the *activiti explorer*. Also Activiti Explorer provides a graphical user interface that exposes functionalities provided by the prototype system. At the data layer of the Activiti is a relational database, which stores business process definitions. We have created a separate relational database to store *annotations* of process models. While Activiti uses the model repository persistence within the *activiti engine* as the persistence solution to access the process definition database, we have introduced and created a separate *process annotation persistence* solution to access the process annotation database.

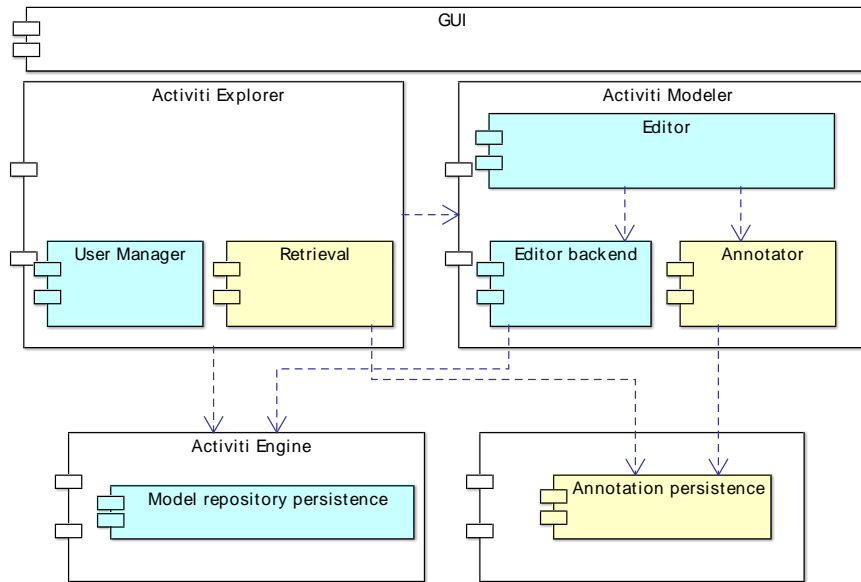


Figure 26: Architecture of a prototype system.

In the following subsections we describe the main components of interest as shown in Figure 26.

6.4.1 Process Modelling Environment

In this prototype, process modelling is provided by Activiti Modeler. Activiti Modeler is a business process modelling framework bringing Web 2.0 technologies to analysts and designers (Rademakers, 2012; Team, 2010). It allows business analysts to model a BPMN 2.0-compliant business process in a Web browser. This implies no client software is needed before one can start modelling. Created process files are stored by the server in a database model repository where each process model is identified by a URL. This means that process models can easily be shared by passing references.

Activiti Modeler is structured into client and server components. The editor, which realizes the process modelling functionality, is the client component. It is a JavaScript application that runs in a Web browser. The editor is based on the Ext-JS, a JavaScript application framework for building Web applications. The backend is the RESTful Web application based on Restlet – a lightweight, comprehensive, open-source API framework for the Java platform (Louvel & Boileau, 2009). It includes a model converter that performs process model conversion from JSON to BPMN format and vice versa. It also provides APIs to Activiti Engine, which includes a model persistence repository that enables access to the process model database when models are created or are to be accessed. That backend also defines and stores stencil sets of process modelling language, in this case BPMN. A stencil set defines explicit typing of the model elements, connection rules, and the visual appearance of elements (Decker et al, 2008).

There are various ways in which Activiti Modeler can be extended. BPMN elements are described with JSON and references to the SVG file. This allows the addition of new BPMN elements or change attributes. Also new modelling languages can be supported by adding stencil sets. Furthermore, features extensions via plugins provided the Activiti Modeler allow to add new functionality and the use of Ext-JS and Restlet make it possible to be customized (Decker et al, 2008). Therefore new Web services can be added.

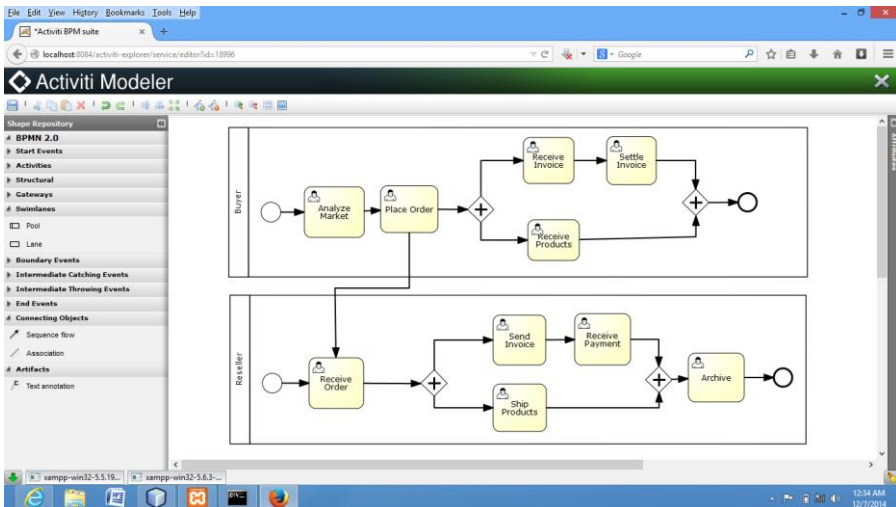


Figure 27: Screenshot for creating and storing a new process model.



Figure 28: Screenshot for annotating a process model.

6.4.2 Semantic Annotation Service

The semantic annotation component is built as an extension of Activiti Modeler. Like the modeller the annotation service is logically structured into client-side and server-side components. The client-side component is a JavaScript application based on the Ext-JS framework. The client side renders the user interface of the annotation service through which a user annotates a process model. The server-side components are three RESTful Web services. One is the Web resource, which is invoked once the model is saved and before the annotation interface is rendered. It extracts the meta info of the process model from the Activiti database and presents the results with the annotation form interface that is rendered to the user to complete the task of annotating the process model. Another Web service resource extracts the annotation of the process from the process annotation database. It returns default annotation values for a new process model, and specific annotation values for the existing process model. The third Web service defines annotation elements and possible values of each element that makes up the annotation model. It also provides an API that accesses the annotation persistence that is the JPAcontainer.

A process model is created using the editor, a component of Activiti Modeler, and then saved into the process model database. Before the model, in JSON, is saved it is converted into BPMN format and stored in a database. Once the model is saved, the annotation service is triggered. The annotation service extracts basic process annotation including business process name, process description and the ID, which point to the address where the model is stored. It then renders the annotation interface (see Figure 28) with the initial annotation of the process model. The interface allows the user to complete the annotation of process and produces process annotation, which is stored in the process annotation database.

6.4.3 Advanced Process Retrieval

The retrieval component, which consists of search and navigation services, is built as an extension of the Activiti Explorer. The Activiti Explorer is a Vaadin application that follows a three-layered architecture. It is organized into a presentation layer, which provides the overall graphical user interface (GUI), the domain model layer, which defines the data model, and the data store layer. At the presentation layer of the Activiti Explorer we have customized the graphical user interfaces by extending GUI components such as the menu bar and view manager. We have also created search and navigation interfaces, which renders a search interface for searching process models and the search results. At the domain model there is a class that defines the annotation elements and associated values as the data model. At

the data store there is a process annotation database, which stores process annotation of stored process models.

The retrieval component uses process annotation persistence, which is the JPAContainer, to access the annotation database.

The search service is designed to be more efficient and effective. This is achieved by allowing users to retrieve all process models that fully or partially match the specifications of the requested process annotation. Permitting the retrieval of partial matches would increase the recall – the proportion of the number of relevant process models retrieved to the total number of relevant process models in the repository (Ali & Du, 2004). Also the navigation structure is implemented based on the annotation model (CPSAM) (Elias & Johannesson, 2013). Therefore, the process models stored in the repository are organized by elements of the annotation model.

Figure 29 presents the annotation-based search, which consists of two parts. The first part is the search interface (the left part of the screen), which allow users to establish a search query in the repository. The second part is the search results interface (the right part of the screen), which displays a list of process models returned from the established search. By limiting the search to annotation elements a user can efficiently retrieve relevant process models from the repository. For example, a user looking for an order-to-cash process model, for delivery of products such as personal computers, may start with *process area*, *process type*, *process phase*, *organization level*, *principal actor* and *other actor*, *resource received* and *resource provided* to form a query. The result of this query is a list of two process models (Figure 22 (a) and (b) shown in Section 5.5) as versions of an order-to-cash process. Thereafter, a user may wish to proceed to view the process, by clicking the process name “Order to cash”, which will display process annotation as shown in Figure 30. Otherwise the user may wish to view a process model by clicking “View process model” from the search results (Figure 29) or from the displayed process annotation (Figure 30). The view process model will retrieve the process model from the process model database and trigger the Activiti Modeler through which the process model is displayed. The user may proceed to export the process model, or edit and store it in the repository.

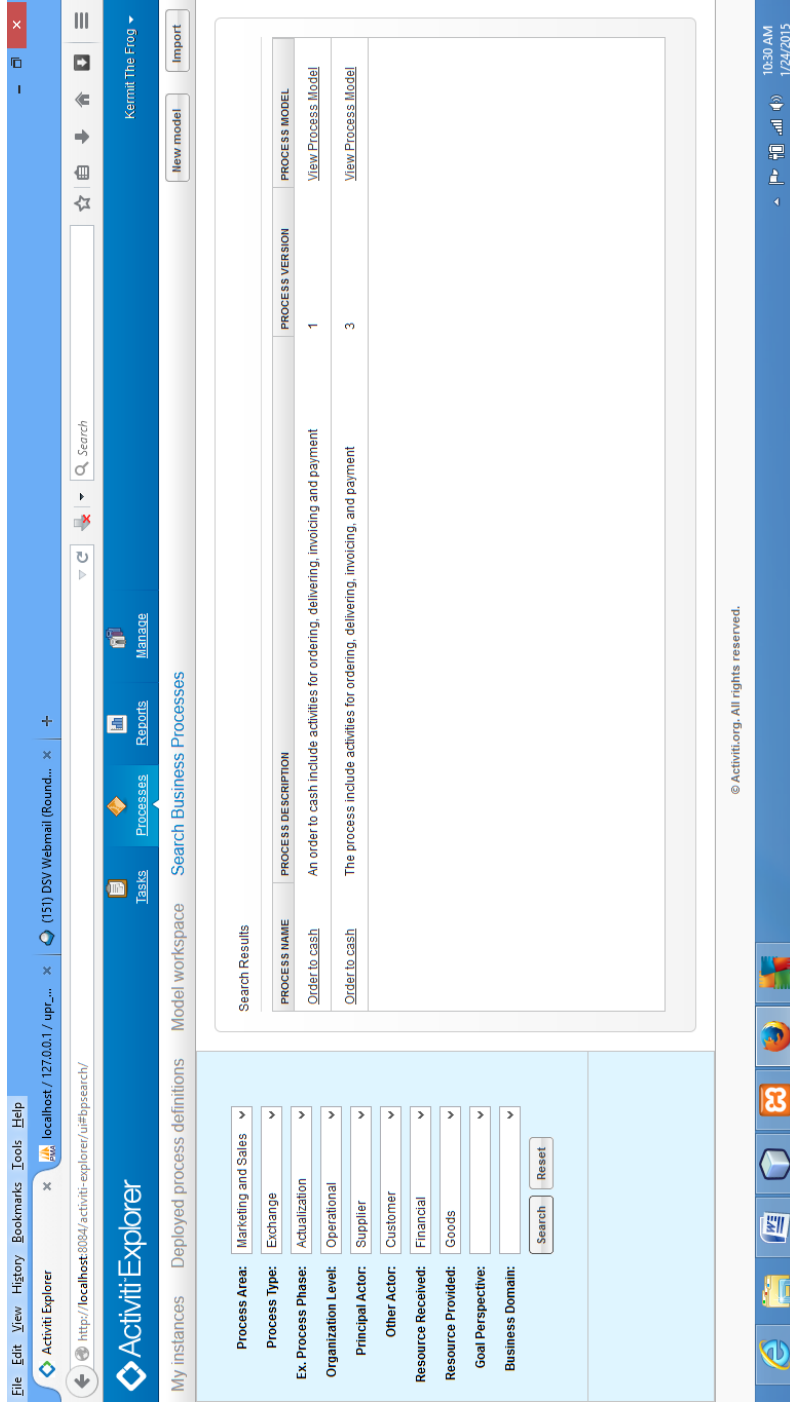


Figure 29: Searching a process model from the repository.



Figure 30: Process annotation of process model.

6.5 Summary and Discussion

The semantic annotation tool implements two main components based on the semantic annotation model (CPSAM). The first component is a semantic annotation service for annotating process models, as they are stored in the repository. The second component is an advanced retrieval service for efficient and effective retrieval of process models from the repository. The annotation service provides a way to annotate process models with the values or terms of annotation elements that make the context-based process semantic annotation model (CPSAM).

The primary goal of the CPSAM model is to improve the efficiency and effectiveness of the process retrieval from the repository. Therefore the advanced retrieval service implements a process model retrieval system, which allows users to direct their search and navigation by using elements of the annotation model.

Process models are created using Activiti Modeler, a process modelling tool, and are annotated, as they are stored in the process database. The annotation service lets users annotate process models and stores the annotation in the process annotation database. During process model retrieval, the retrieval component of the prototype tool enables users to retrieve all relevant process models that are fully or partially matched to the requested process annotation specifications. This is how the semantic annotation model improves the efficiency and effectiveness of model retrieval.

An evaluation of the semantic annotation model to test whether it improves the efficiency and effectiveness of the retrieval system is presented in the next chapter.

7 Evaluation of the Semantic Annotation Model

7.1 Overview

This chapter presents and discusses the process and results of two studies we carried out to empirically evaluate CPSAM. The purpose of the first study was to evaluate CPSAM from the annotator's point of view, meaning that annotation consistency, annotation correctness and perceived ease of annotation were investigated. The purpose of the second study was to evaluate CPSAM from the user's point of view, meaning that the perceived ease of use, repository search and navigation, and understandability were investigated. As part of the second experiment, we tested the discriminatory power of CPSAM. This chapter builds on Paper IV and V as listed in section 1.5 of Chapter 1. The chapter begins by describing the evaluation framework we have used to evaluate the annotation model. It goes on to separately discuss the two studies carried out through controlled experiments. For each study we describe experimental settings, studied variables, results and discussions. As part of the experimental settings we discuss how participants were selected, preparation of the experiments, which includes experimental materials, and the conduction of the experiments. For the studied variables we define and distinguish two types of variables: performance- and perception-based variables in accordance with the evaluation framework. For the results and discussion we present data collected from the experiments, analysis and interpretation of the data. The chapter then present a summary of the annotation model evaluation based on the two studies according to the requirements as defined in Chapter 5. Finally the chapter concludes with some key remarks about the annotation model evaluation.

7.2 The Evaluation Framework

Our evaluation framework follows the Method Evaluation Model (MEM) (Moody, 2003), a method for evaluating IS design methods. The MEM is based on two areas of theory: the Technology Acceptance Model (TAM) (Davis et al, 1989), from the IS success literature, and Methodological Pragmatism (Rescher, 1997), from the philosophy of science. The MEM is chosen because it incorporates both aspects of evaluation, i.e. performance and user perception. These are strong and quite desirable features of the

MEM model. In addition, the MEM model includes aspects for measuring the behaviour of the users by capturing the actual usage of the model. The approach is quite relevant to evaluating the semantic annotation model since one would be interested in measuring (i) the success of the annotating effort in addition to measuring the behaviour of the annotators, and (ii) the success of using the annotation model for searching, navigation and understanding process models.

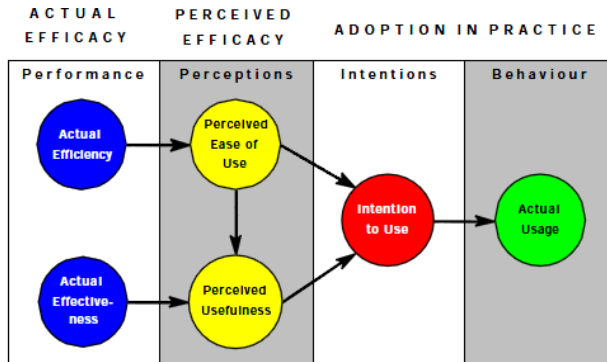


Figure 31: The Method Evaluation Model (MEM) (Moody, 2003).

Figure 31 shows the MEM indicating the primary constructs and causal relationships between them. In the MEM model, actual efficiency measures the extent to which the method is required to perform the act. Whereas actual effectiveness measures the extent to which the method improves the quality of the result. They both determine whether the model improves task performance whereas perceived ease of use and perceived usefulness represent the user’s perceptions about the model’s efficiency and effectiveness. The first three central constructs in the MEM model (performance, perceptions and intentions) represent internal and psychological variables while the last construct represents behavioural constructs that can be measured objectively.

7.2.1.1 Evaluation from Annotator’s Point of View

From the annotator’s point of view the construct of the MEM in the context of applying CPSAM to annotate process models can be defined as follows:

Actual efficacy is the degree to which the use of CPSAM to annotate process models achieves its objectives, which in this case are correctness and consistency of the annotation produced.

Actual efficiency can be measured by investigating the extent to which the annotation of process models using CPSAM is resource efficient (does not require much time and effort). An annotator does not require much effort or time to annotate a process model.

Actual effectiveness can be measured by investigating the extent to which annotating process models using CPSAM is error free. The produced process annotation is error free (i.e. accuracy).

Perceived ease of use is the degree to which a person believes that annotating process models using the CPSAM is free of effort.

7.2.1.2 Evaluation from User's Point of View

From the user's point of view the construct of the MEM in the context of CPSAM can be defined as follows:

Actual efficacy is the degree to which the use of CPSAM-based annotation achieves its objectives, which in this case are searching process models, navigating the process model repository and understanding process models.

Actual efficiency can be measured by investigating the extent to which the use of CPSAM-based annotation reduces the time or effort required to search and navigate the repository to find a relevant process model, as well as the effort or time required to comprehend a process model.

Actual effectiveness can be measured by investigating the extent to which the use of CPSAM-based annotation improves the quality (i.e. accuracy) of the search, navigation and comprehension of process models.

Perceived ease of use is the degree to which a person believes that using the CPSAM-based annotation in searching, navigating and comprehending process models is free of effort.

Perceived usefulness is the degree to which a person believes that the use of CPSAM-based annotation will be effective (i.e. accurate) in searching, navigating and comprehending process models in the repository.

Intention to use is the extent to which a person intends to use CPSAM-based annotation for searching, navigating and comprehending a process model.

7.3 Study I – Validation of Annotation Correctness and Consistency

In this section we describe the study (Elias et al, 2010) we have carried out for empirically evaluating the CPSAM from the annotator's point of view. Specifically, the purpose of the study is to evaluate the consistency and correctness of annotating business processes using CPSAM. Furthermore, the user perception of the model is tested.

7.3.1 Experimental Settings

7.3.1.1 Selecting Participants

The participants involved in the experiment were a mix of master's students in Engineering and Management of Information Systems (EMIS) and PhD

students in Information Systems at the Royal Institute of Technology (KTH). By the time the experiment was done, all students had completed a course on Enterprise Systems and Modelling, in which they learnt basic concepts about business process modelling. The benefit of using student participants is that they form a homogeneous group with respect to their academic background and industrial experience. Furthermore, the experimental tasks did not require a high level of industrial experience, which justifies our selection of the participants.

7.3.1.2 Preparing the Experiment

For the experiment, the following materials were prepared:

- A document defining the CPSAM model and a description of each element (as presented in chapter 5).
- A document describing (five) business processes. In order to increase understanding, processes were presented in both textual and graphical form. Annotating business processes is a time-consuming task, therefore to keep the participants positive toward the experiment, we had to limit the number of processes to five. The decision to limit the business processes was also based on our experience from the pilot study described below.
- A template for annotating business processes. This is a two-dimensional table in which rows present elements of the CPSAM, and columns present the processes to be annotated.
- A post-task survey questionnaire to measure user perception of the model on a scale of 1 to 5 (Strongly Disagree, Disagree, Not Sure, Agree and Strongly Agree).

As part of the preparation for the experiment, a pilot study was conducted with three participants (PhD students). The purpose of the pilot study was to evaluate how well the participants were able to perform the experiment. The results and comments from this study were used to improve the CPSAM element definitions, business process descriptions and the template for the experiment.

7.3.1.3 Conducting the Experiment

In this study (Elias et al, 2010), 30 participants performed a controlled experiment. Participants were given experimental materials, followed by an explanation of the annotation model, how to use the model and a template for annotating processes. Participants were then asked to annotate the business processes without any time constraint. After annotating the business processes, participants completed a post-task survey. Responses were received from 20 participants, making the response rate 66.7%.

7.3.2 Studied Variables

In order to evaluate the consistency and correctness of annotating business processes using CPSAM, and the user perception of the model, the following three variables were defined:

Variable 1. Annotation Consistency (AC): The degree to which process annotation (using CPSAM) by different people is identical. AC is measured by the percentage of participants with identical process annotation in individual elements of CPSAM.

The steps taken for measuring AC are to let different participants annotate a set of business processes and then we computed AC as follows:

1. Let $\text{Max}_{e,p}$ be the maximum number of participants with identical annotations on element e for process p . e is an element of $\{\text{Resource, Actor, Process Level, Process Relationship, Process Area, Process Phase, Process Type}\}$.

For example, suppose a process ($p=1$) is annotated by 20 participants and for an element (*Process Level*), out of the 20 participants 12 annotate it as “operational”, 5 as “tactical” and 3 as “strategic”. Therefore, $\text{Max}_{\text{Processlevel},1}=12$.

2. Annotation Consistency on element e for process p , $\text{AC}_{e,p} = (\text{Max}_{e,p} * 100)/N$, where N is the total number of participants. For the example given above, $\text{AC}_{\text{Processlevel},1} = (12 * 100)/20 = 60$.
3. The average AC for an element e , $\text{AC}_e = (\sum \text{AC}_{e,p}) / n$ for $p = 1 \dots n$, where n is the number of annotated processes.

The existence of similarities in process annotation means that there is a common understanding of the CPSAM between different people. This implies that the process annotations based on CPSAM will communicate the same meaning to different people.

Variable 2. Annotation Correctness (AR): The degree to which process annotation (using CPSAM) by different people is correct. AR is measured by the percentage of participants who correctly annotated a process for an element of CPSAM.

The majority of participants may have a common but incorrect understanding of the CPSAM model. Therefore, in order to determine whether the process annotation by participants is correct or not, the AR is measured.

For measuring the AR, the process annotation from different participants is compared with the process annotation from the inventors of CPSAM, assuming that the inventors' annotation is correct. AR is computed as follows:

1. Let $C_{e,p}$ be the number of participants with correct (identical to inventors') annotation on element e for process p , where e is an element

in {*Resource, Actor, Process Level, Process Relationship, Process Area, Process Phase, Process Type*}.

For example, suppose a process ($p=I$) is annotated by 20 participants and for an element (*Process Level*), out of 20 participants 12 annotate it as “operational”, 5 as “tactical” and 3 as “strategic”, where the correct (inventors’) annotation is “tactical”. Therefore, $C_{Processlevel,I}=5$.

2. The Annotation Correctness on element e for process p $AR_{e,p} = (C_{e,p} * 100)/N$, where N is the total number of participants. For the example given above $AR_{Processlevel,I} = (5 * 100)/20 = 25$. Similarly, if the correct (inventors’) annotation is “operational” then $AR_{Processlevel,I} = 60$.
3. The average AR for an element e , $AR_e = (\sum AR_{e,p}) / n$ for $p = 1 \dots n$, where n is the number of annotated processes.

The existence of similarities in process annotation (between participants and inventors) means that the CPSAM model elements are correctly understood. This implies that the process annotation produced will be free of errors.

Variable 3. Perceived Ease of Use (PEOU): The degree to which a person believes that using CPSAM for annotating processes would be free of effort. In order to investigate perceived ease of use we asked the participants to assess two statements, shown in Table 9, on a scale of 1 to 5 (Strongly Disagree to Strongly Agree).

Table 9: Items for measuring perception-based variables

Items	Statements
PEOU1	The annotation definitions are clear and helpful for annotation
PEOU2	It was easy to annotate the business processes

7.3.3 Results and Discussion

In this section, the data collected from the experiment are analysed and discussed in order to evaluate the CPSAM element definitions. For the analysis, the mean and the standard deviation of annotation consistency and correctness for each element are computed. Table 10 and Table 11 and Figure 32 show a summary of statistics of process annotation.

7.3.3.1 Annotation Consistency (AC)

The results (in Figure 32) show that more than 62% of participants have identical process annotation for the following elements: *Resource, Actor, Process Level, Process Relationship* and *Process Type*. This indicates that there is a common understanding of these CPSAM elements between different users. However, the *Process Area* and *Process Phase* elements have less than 50% of participants with identical process annotation. This indicates that the *Process Area* and *Process Phase* definitions are differently understood by the participants.

Table 10: Annotation consistency

	Resource	Actor	Process Level	Process Relationship	Process Area	Process Phase	Process Type
Mean	86.56	76.46	69.58	80.56	44.62	48.69	62.24
StDIV	12.40	12.24	21.19	9.62	3.23	14.16	14.60

Table 11: Annotation correctness

	Resource	Actor	Process Level	Process Relationship	Process Area	Process Phase	Process Type
Mean	86.56	76.46	69.58	80.56	36.84	42.64	54.66
StDIV	12.40	12.24	21.19	9.62	9.21	17.86	20.06

7.3.3.2 Annotation Correctness (AR)

The results (in Figure 32) show that more than 54% of the participants have correctly annotated business processes for the elements *Resource*, *Actor*, *Process Level*, *Process Relationship* and *Process Type*. This indicates that the CPSAM element definitions are well understood by different people, implying that most process annotations generated by users based on the model will be free of errors. However, less than 50% of participants have correctly annotated the elements *Process Area* and *Process Phase*. The detailed analysis shows that participants who correctly annotated *Process Area* and *Process Phase* had more industrial experience than the others. While the two elements are based on widely accepted frameworks (McCarthy, 1982; Dunn et al., 2005), understanding and applying these definitions to annotating business processes seems to require some basic industrial experience, which some participants lacked.

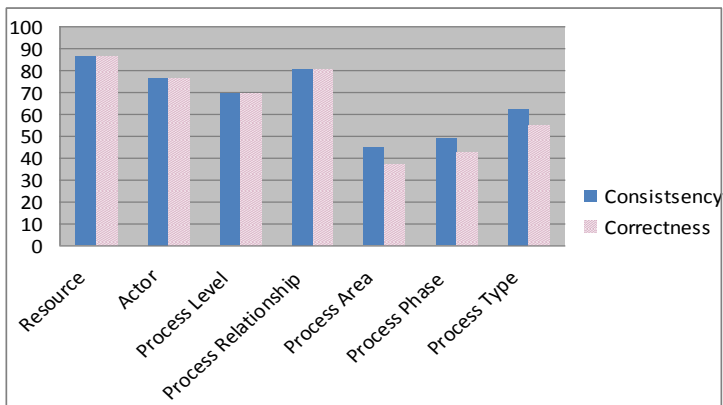


Figure 32: Annotation consistency and correctness.

Comparing the annotation consistency (AC) and correctness (AR), Figure 32 shows that AC is equal to AR for the elements *Resource*, *Actor*, *Process Level* and *Process Relationship*. This means that, for these elements, the majority of the participants who identically annotated the processes were

correct. Therefore, the majority of the participants have a common and correct understanding of CPSAM and its element definitions. However, the definitions of Process Area, Process Phase and Process Type were not correctly annotated, so we hypothesize that these definitions need to be sharpened.

7.3.3.3 Perceived Ease of Annotation (PEOU)

Figure 33 shows the summary of statistics for user perception of the model. More than 52% of the participants agree (agree and strongly agree) with PEOU1 (*The annotation definitions are clear and helpful for annotation*). More than 58% of the participants agree (agree and strongly agree) with PEOU2 (*It was easy to annotate the business processes*).

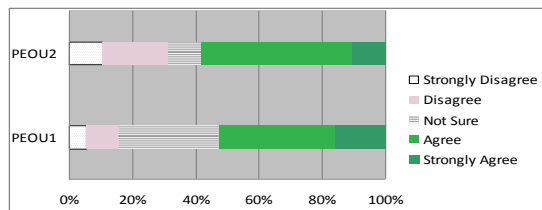


Figure 33: Perceived ease of annotation.

7.4 Study II – Validation of User Performance and Perception

In this section we describe the study (Elias & Johannesson, 2012a) we have carried out to empirically evaluate the CPSAM model from the user’s perspective. Specifically, the purpose of this study is to test whether the CPSAM meets its objectives – to facilitate searching of process models, to support navigation of the repository and enhance understandability of process models. In this research we have carried out a two-stage evaluation test. In the first stage, we used a controlled experiment to test the performance (actual efficacy). In the second stage, a post-task survey questionnaire was used to test the user’s perception of the annotation model.

7.4.1 Experimental Settings

7.4.1.1 Hypothesis

For this purpose, we formulated the following hypotheses for measuring the performance of the annotation model in searching, navigation and understandability of process models.

H1: The annotation positively affects searching of process models in the repository

H2: The annotation positively affects navigation of the process model repository

H3: The annotation positively affects process model understandability

In the second stage, we used a survey (using a questionnaire) to test the user's perception of the effect of annotation. In the following subsections, a detailed description of the experimental design is presented.

7.4.1.2 Experimental Materials and Tasks

The main instrumentation for the experiment was a repository prototype demonstrated in Chapter 6. For the experiment, the repository was populated with more than 100 business process models adopted from existing repositories, i.e. MIT, IBM and SAP. The process models were redesigned using BPMN, a standard process modelling notation, and stored in the repository. In addition to that, the following materials were prepared for the experiment:

- A document describing the CPSAM annotation model
- A document describing the prototype of the repository
- A post-task survey questionnaire to measure the user perception of the effect of the annotation. The survey consisted of eight closed questions assessed over a scale of 1 to 5 (Strongly Disagree, Disagree, Not Sure, Agree and Strongly Agree)
- A document describing a set of questions for each experimental task. The document included five questions for searching, five questions for navigation and four questions for understanding, i.e. one question for each process design perspective (functional, behavioural, organizational and informational).

The questions related to the understanding task were accompanied by six process models (P1 to P6), where half of the accompanied processes were annotated and the others were not. During the experiment the participants were divided into two groups. The first group of participants was given P1, P3 and P5 as annotated process models, whereas the second group was given P2, P4 and P6 as annotated process models.

The experiment consisted of four main tasks, which are as follows:

- *Searching Task.* In this task participants were asked to find process models that are relevant to a given question. For each question, participants were required to perform both a keyword-based and an annotation-based search. From the questions, participants were supposed to identify some keywords and annotations that they would use for searching.
- *Navigation Task.* In this task participants were asked to navigate the repository and locate process models that are relevant to a given

question. For each question, participants were required to perform both alphabetical-based and annotation-based navigation. From the questions, participants were supposed to identify some alphabets of the keyword and annotations that would guide them in navigating the repository.

- *Understandability Task*. In this task participants were required to study process models and answer related questions. As discussed above, half of the process models were annotated and the other half were unannotated.
- *Post-Task Survey*. Upon completing the experiment participants were asked to perform a post-task survey.

7.4.1.3 Participant Selection and Experimental Treatment

The participants involved in the experiment were a mix of master's students in Engineering and Management of Information Systems (EMIS) and PhD students in Information Systems at KTH. By the time the experiment was done, all students had completed a course on Enterprise Systems and Modelling, in which they learnt basic concepts about business process modelling. The benefit of using student participants is that they form a homogeneous group with respect to their academic background and industrial experience. Furthermore, the experimental tasks did not require high levels of industrial experience, which justifies our selection of the participants.

At the beginning of the experiment, the participants were given a short list of written instructions describing the experiment. Experiment mentors demonstrated how the prototype could be used to search and navigate the repository. Furthermore, a case of understanding process model annotation was demonstrated.

For the experiment, 20 randomly selected participants were given the materials (described above). Responses from 15 participants were received and all the collected data were considered for analysis. Due to the length of the experiment each participant was asked to perform two tasks and a post-task survey. For the analysis (in Section 7.4.3) of each task (searching, navigation and understandability) results from 10 participants are included.

7.4.2 Studied Variables

In order to test the influence of the annotations (based on the CPSAM) on searching, navigation and process model understandability, we distinguish two types of variable: performance-based (objective) and perception-based (subjective) measures.

7.4.2.1 Performance-Based Variables

Variable 1. Search Correctness (SC): The degree of accuracy with which a user finds a relevant process model by searching the repository. It is measured in terms of F-measure – the harmonic mean of *precision* and

recall. *Precision* is the fraction of retrieved process models that are relevant, whereas *recall* is the fraction of relevant process models retrieved. F-measure is a standard measure for evaluating information retrieval results (Kandefor & Shapiro, 2009).

The steps taken for measuring *SC* are to let different participants find a process model(s) from the repository that is relevant to a given question and then we compute *SC* as follows:

1. Let $RR(q, i)$ be the total number of relevant processes retrieved on question q by subject i , $IR(q, i)$ be the number of irrelevant processes retrieved on question q by subject i , and $RN(q, i)$ be the number of relevant processes in the repository that have not been retrieved. The precision $PR(q, i)$ and recall $RC(q, i)$ on question q by subject i are $PR(q, i) = RR(q, i) / [RR(q, i) + IR(q, i)]$ and $RC(q, i) = RR(q, i) / [RR(q, i) + RN(q, i)]$.
2. The Search Correctness on question q by subject i , measured by F-measure, $F(q, i) = 2 * PR(q, i) * RC(q, i) / [PR(q, i) + RC(q, i)]$.
3. The average *SC* by subject i , $F(i) = (\sum F(q, i)) / n$ for $q = 1$ to $q = n$, where n is the number of process retrieval questions.

By comparing the results of *SC* for the keyword-based and annotation-based search, we can determine whether the annotation positively affects the searching or not.

Variable 2. Navigation Efficiency (NE): Is the proportion of the steps (efforts) that are useful to find the relevant process models in the repository. It is measured by the minimum path length (*MPL*) divided by the total user path length (*TUPL*) (Huang, 2003; Mondary et al, 2007) used to locate the process model. The path length is the number of steps (button clicks) performed in order to find relevant process models by navigating the repository. The total user path length is the total number of steps a user used to locate a relevant process model by navigating the repository. The minimum path length is the least number of steps needed to locate a relevant process model.

The steps taken for measuring *NE* are to let participants locate a process model relevant to a given question by navigating the repository and then we compute *NE* as follows:

1. Let $MPL(q)$ be the least number of steps needed to locate a process model for question q by navigating the repository and $TUPL(q, i)$ be the total number of steps subject i used to locate relevant processes for question q .
2. The Navigation Efficiency for locating a process for question q by subject i is $NE(q, i) = MPL(q) / TUPL(q, i)$.
3. The average *NE* by subject i , $NE(i) = (\sum NE(q, i)) / n$ for $q = 1$ to $q = n$, where n is the number of process retrieval questions.

By comparing the results of *NE* for alphabetic-based and annotation-based navigation, we can determine whether the annotation positively affects the navigation or not.

Variable 3. Understandability (*UL*): This is the degree of correctness to which a user understands a process model. It is measured as the fraction of correct answers given by the subject to the different questions about the process (Melcher et al, 2010).

The steps taken for measuring *UL* are to let subjects study a process model and respond to questions related to the process model. We then compute *UL* as follows:

1. Let $CA(p, i)$ be the number of correct answers on process p by subject i and $EA(p)$ be the number of expected correct answers on process p . The understandability on process p by subject i is given by $UL(p, i) = CA(p, i) / EA(p)$.
2. The average understandability by subject i is $UL(i) = (\sum UL(p, i)) / n$ for $p=1$ to $p=n$, where n is the number of process models.

By comparing the results of *UL* for annotated and unannotated process models, we can determine whether the annotation positively affects the understandability or not.

Table 12: Items for measuring perception-based variables

<i>Items</i>	<i>Statements</i>
PU1	I think the annotations have improved the process of locating and searching process models in the repository
PU2	I found navigating the process model repository based on CPSAM elements to have improved my work
PU3	I found the annotations to be helpful for understanding process models
PEOU1	It was easy for me to locate/search the process models
PEOU2	It was easy for me to navigate the repository
IU1	If I have to search a process in the repository in the future I will use annotation-based search
IU2	If I have to navigate the repository in the future I will use an annotation-based approach
IU3	If I am involved in building the repository for process models I will recommend the CPSAM model

7.4.2.2 Perception-Based Variables

Variable 4. Perceived Usefulness (*PU*): Is the degree to which a person believes that the annotations (based on CPSAM) improve searching, navigation and understanding of process models.

Variable 5. Perceived Ease of Use (PEOU): Is the degree to which a person believes that the use of the CPSAM-based annotation is free of effort.

Variable 6. Intention to Use (IU): Is the extent to which a person intends to use the CPSAM-based annotation for searching, navigating and comprehending a process.

In order to investigate users' perception of the model we asked the participants to assess several statements on a scale of 1 to 5 (Strongly Disagree, Disagree, Not Sure, Agree and Strongly Agree). The statements (as shown in Table 12) are PU1, PU2 and PU3 for *Perceived Usefulness*, PEOU1 and PEOU2 for *Perceived Ease of Use* and IU1, IU2 and IU3 for *Intention to Use*.

7.4.3 Results and Discussion

In this section, the data collected from the study are analysed and discussed in order to evaluate the CPSAM. The Wilcoxon matched-pairs test is chosen for data analysis, as demonstrated by Zobel (1998); Wilcoxon's signed-rank test is a reliable way to evaluate statistical differences between two retrieval systems. The following are the results and discussion for each studied variable.

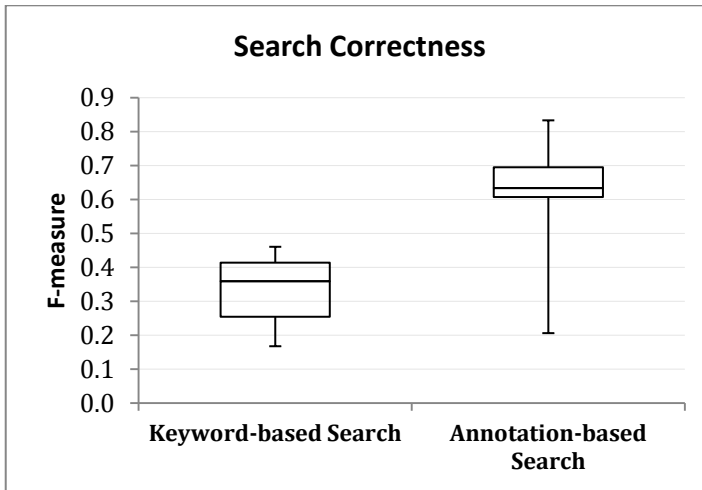
7.4.3.1 Effects of Annotation on Searching of Process Models

Figure 34 depicts box plots (a) and a graph (b) of the F-measure for keyword-based search and annotation-based search. The plot and the results in Table 13 show that Search Correctness for annotation-based search (ABS) is better than keyword-based search (KBS) given that the median F-measure value is higher for ABS than for KBS. This indicates that annotation positively affects the searching of process models in the repository.

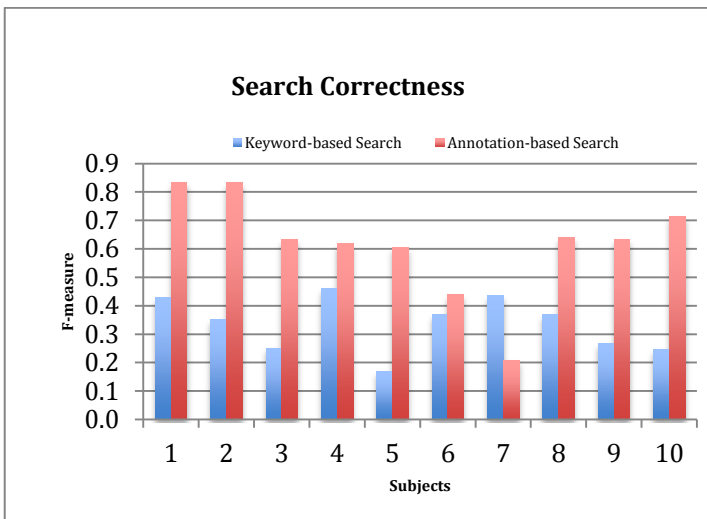
Using the data in Table 13 and the Wilcoxon signed-rank test and a 0.01 significance level, we test the claim that there is no difference between keyword-based search (KBS) and annotation-based search (ABS). The sum of the absolute values of the negative ranks $T^- = 52$ and the sum of the

Table 13: Search correctness

Subjects	F-KBS	F-ABS	Differences	Diff. Rank	Signed Rank
1	0.430000000	0.833333333	-0.403333333	7	-7
2	0.351052632	0.833333333	-0.482280702	10	-10
3	0.251052632	0.633333333	-0.382280702	6	-6
4	0.460769231	0.617142857	-0.156373626	2	-2
5	0.167719298	0.604617605	-0.436898306	8	-8
6	0.367719298	0.440000000	-0.072280702	1	-1
7	0.434385965	0.206349206	0.228036759	3	3
8	0.367719298	0.638961039	-0.271241741	4	-4
9	0.266666667	0.633333333	-0.366666667	5	-5
10	0.246000000	0.713333333	-0.467333333	9	-9



(a)



(b)

Figure 34: Search correctness.

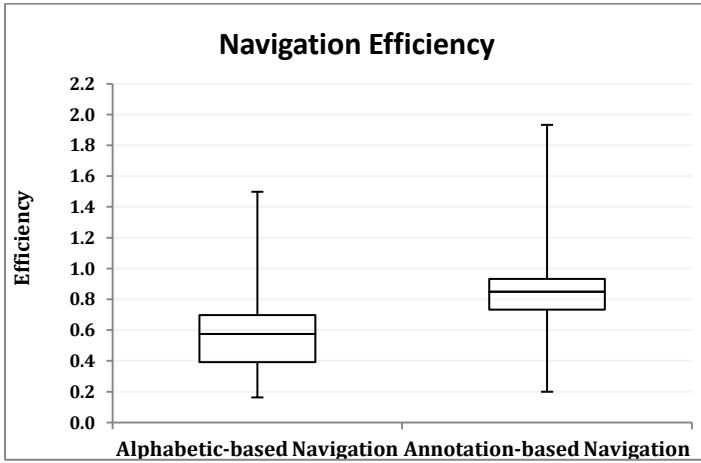
positive ranks $T^+=3$. Because $n=10$, we have $n \leq 30$, so we use a test statistic of $T=3$. Therefore, Critical Value $(0.01(2), 10)=3.0$ and Critical Value $(0.01(1), 10)=5.0$. Since T^+ is less than or equal to the Critical Value we reject the null hypothesis. Thus hypothesis H_1 is accepted.

7.4.3.2 Effects of Annotation on Navigating the Process Model Repository

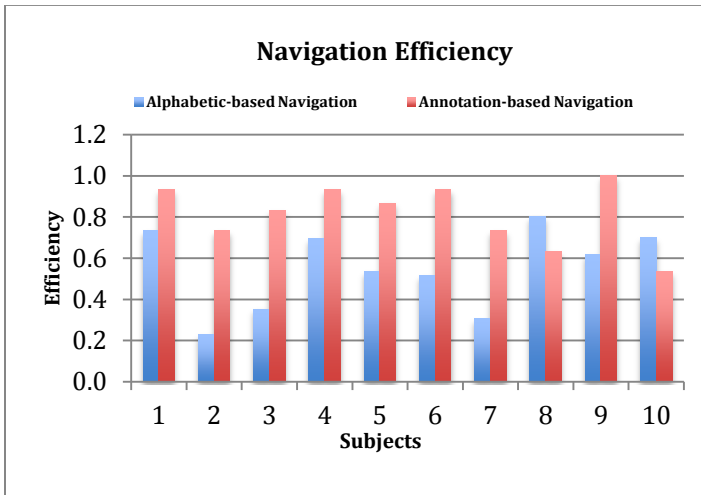
Figure 35 depicts box plots (a) and a graph (b) of navigation efficiency for alphabetic-based navigation (ALN) and annotation-based navigation (ANN). The plot and the results in Table 14 show that Navigation Efficiency (NE) for annotation-based navigation (ANN) is better than alphabetic-based

Table 14: Navigation Efficiency

Subjects	NE-ALB	NE-ANB	Differences	Diff. Rank	Signed Rank
1	0.733333333	0.933333333	-0.200000000	3	-3
2	0.230000000	0.733333333	-0.503333333	10	-10
3	0.351904762	0.833333333	-0.481428571	9	-9
4	0.695238095	0.933333333	-0.238095238	4	-4
5	0.533333333	0.866666667	-0.333333333	5	-5
6	0.516666667	0.933333333	-0.416666667	7	-7
7	0.306666667	0.733333333	-0.426666667	8	-8
8	0.800000000	0.633333333	0.166666667	1.5	1.5
9	0.616666667	1.000000000	-0.383333333	6	-6
10	0.700000000	0.533333333	0.166666667	1.5	1.5



(a)



(b)

Figure 35: Navigation efficiency.

navigation (ALN) given that the median value of *NE* is higher for ANN than for ALN. This indicates that annotation positively affects the navigation performance in the repository.

Using the data in Table 14 and the Wilcoxon signed-rank test and a 0.05 significance level, we test the claim that there is no difference between alphabetic-based navigation (ALN) and annotation-based navigation (ANN). The sum (T-) of the absolute values of the negative ranks is 52 and the sum (T+) of the positive ranks is 3. Because $n=10$, we have $n \leq 30$, so we use a test statistic of $T=3$. Therefore, Critical Value $(0.05(2), 10)=8.0$ and Critical Value $(0.05(1), 10)=10$. Since $T+$ is less than or equal to the Critical Value we reject the null hypothesis and accept hypothesis H_2 .

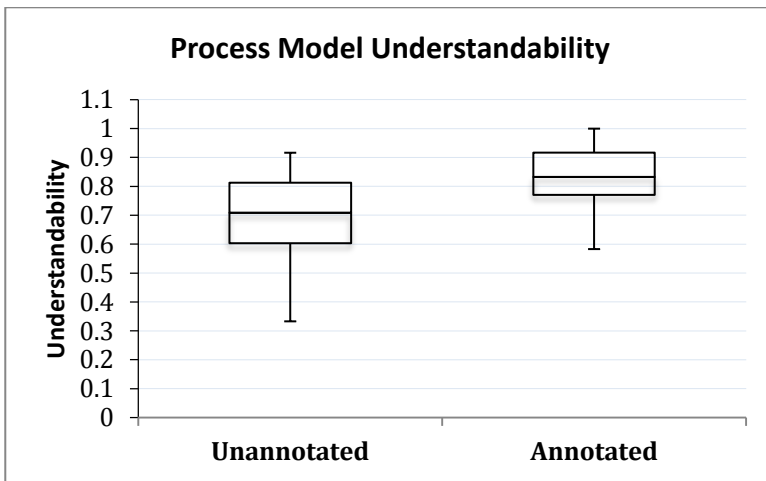
7.4.3.3 Effects of Annotation on Process Model Understandability

Figure 36 depicts box plots (a) and a graph (b) of understandability (UL) of unannotated and annotated process models. The box plots and the graph show that understandability of annotated process models is higher than understandability of unannotated process models, given that the median of understandability for annotated models is higher than the median of understandability of unannotated models. This indicates that annotation positively affects process model understandability.

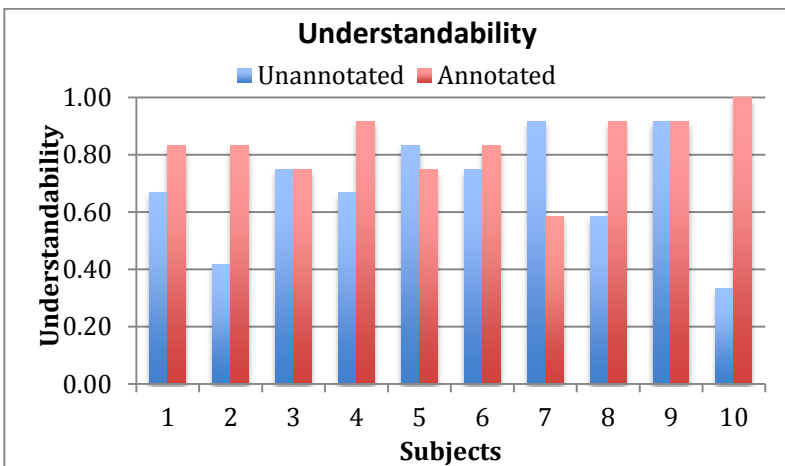
Table 15: Process model understandability

Subjects	Unannotated	Annotated	Differences	Diff. Rank	Signed Rank
1	0.666666667	0.833333333	-0.166666667	3	-3
2	0.416666667	0.833333333	-0.416666667	7	-7
3	0.750000000	0.750000000	0.000000000	-	
4	0.666666667	0.916666667	-0.250000000	4	-4
5	0.833333333	0.750000000	0.083333333	1.5	1.5
6	0.750000000	0.833333333	-0.083333333	1.5	-1.5
7	0.916666667	0.583333333	0.333333333	5.5	-5.5
8	0.583333333	0.916666667	-0.333333333	5.5	-5.5
9	0.916666667	0.916666667	0.000000000	-	
10	0.333333333	1.000000000	-0.666666667	8	-8

Using the collected data and the Wilcoxon signed-rank test and a 0.05 significance level, we test the claim that there is no difference between understandability of unannotated models and annotated models. The sum of the absolute values of the negative ranks $T-= 29$ and positive ranks $T+=7$. Because $n=8$ (we omit two values with difference = 0), we have $n \leq 30$, so we use a test statistic of $T=7$. Therefore, Critical Value $(0.05(2), 8)=3.0$ and Critical Value $(0.05(1), 8)=5.0$. Since neither $T+$ nor $T-$ is less than or equal to the Critical Value we fail to reject the null hypothesis.



(a)



(b)

Figure 36: Process model understandability.

7.4.3.4 User Perception of the Annotation

Figure 37 shows the summary of statistics for user perception of the annotation model.

Perceived Usefulness (PU). The graph shows that more than 80% of participants at least agree (i.e. agree and strongly agree) with PU2 and PU3, whereas more than 60% at least agree on PU1. Therefore, it can be argued that most users perceived the annotation to be useful.

Perceived Ease of Use (PEOU). The graph shows that more than 50% at least agree on PEOU1, whereas more than 80% of participants at least agree with PEOU2. Therefore, it can be argued that most users perceived the annotation to be easy to use.

Intention to Use (IU). The graph shows that less than 30% of participants at least agree with IU1 (i.e. more than 60% of participants are not sure about IU1), whereas more than 60% at least agree with IU2 and IU3. This implies that most users are not sure about their intention to use annotation-based search. We hypothesize that the reasons for this are: a) that annotation-based search requires some time and effort to think about different annotation elements before searching, and b) people are used to keyword-based search as their searching routine.

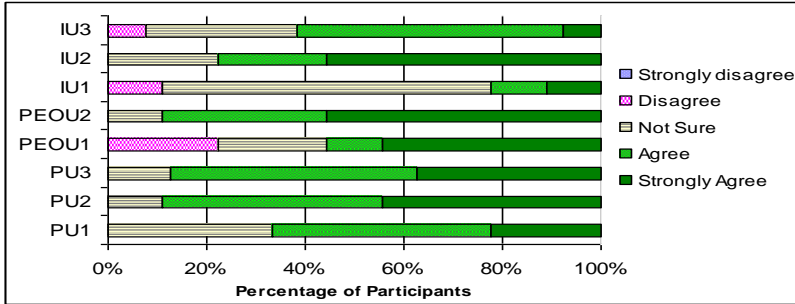


Figure 37: User perception of the annotation model.

7.4.3.5 Annotation Discrimination

To analyse the level of process annotation discrimination, we have randomly selected 100 annotated processes in the repository. Figure 38 depicts the number of processes with the same annotation (vertical axis) in 25, 50, 75 and 100 per cent of the annotation elements (horizontal axis). The result shows that 24, 10, 6 and 6 processes had the same annotation in 25, 50, 75 and 100 per cent of the annotation elements, respectively. The detailed analysis shows that process annotations are well discriminated in all elements, except in “Organization Level”, which has the following annotation: 62 processes annotated as “Operational”, 18 as “Tactical” and 20 as “Strategic”. Thus the study indicates that the discriminatory power is high, as only a small percentage of the process models are annotated in the same way.

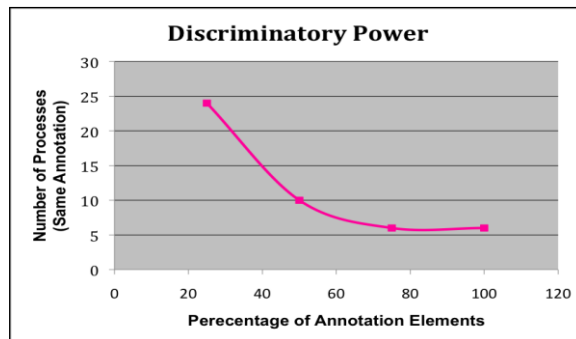


Figure 38: Annotation discriminating power.

7.5 The Evaluation

In this section we present an evaluation of CPSAM according to the requirements of Section 5.2. The evaluation is based on the studies described above as well as informed argument (Hevner et al, 2004; Peffers et al, 2012).

- *High Annotation Consistency.* The results from study I above indicate that the annotation consistency is high for the annotation elements *Resource*, *Actor*, *Process Level*, *Process Relationship* and *Process Type*. However, for two annotation elements, *Process Area* and *Process Phase*, the annotation consistency in study I was low. Thus, there seems to be a need for improved explanations of these elements and possibly also instructions for their use.
- *High Annotation Correctness.* The results from study I indicate that annotation correctness is high for the annotation elements *Resource*, *Actor*, *Process Level*, *Process Relationship* and *Process Type*. However, similarly to the results for annotation consistency, the annotation correctness is lower for *Process Area* and *Process Phase*. Furthermore, study participants who correctly annotated *Process Area* and *Process Phase* had extensive industrial experience. In summary, the elements *Process Area* and *Process Phase* stand out as more challenging than the others and will require further investigation.
- *High Perceived Ease of Annotation.* The results from study I indicate that most users perceive annotation as an easy task, as almost 60% of the participants agree or strongly agree with the statement that it was easy to annotate process models using CPSAM. We believe that this number can be considerably increased by addressing the issues with the *Process Area* and *Process Phase* elements.
- *High Perceived Ease of Use.* The results from study II indicate that users perceive CPSAM as easy to use. Nevertheless, most users do not express an intention to use annotation-based search in place of keyword-based search, which is a challenge to be addressed.
- *High Usability.* The results from study II are ambiguous for this requirement. On the one hand, study II indicates that search correctness and navigation efficiency benefit from an annotation-based approach, while this is not the case for navigation effectiveness. One of the reasons could be that most participants did not use a combination of annotation elements due to the small number of process models currently in the repository.
- *Enhanced Understandability.* The results from study II indicate that annotated process models are easier to understand than unannotated ones, though the difference is not dramatic. It is possible that there would be bigger differences for more complex models than those used in the study.
- *High Discriminatory Power.* The results from study II indicate that process models are not annotated in the same way. However, these

results are preliminary, as they are based on a small number of process models.

- *Extensibility*. Several of the annotation elements can be tailored to the domain under consideration. For example, additional categories can be added to the *Resource* annotation element. The same possibility also applies to the *Actor* and *Business Context* elements.

7.6 Summary and Discussion

CPSAM has been empirically evaluated in two different studies through controlled experiments. In the first study, the model was evaluated to test the consistency and correctness of process annotation. From the study, we have learnt that both the annotation consistency and correctness are high for most of the elements. This implies that the given definitions of CPSAM elements are understandable.

In the second study, we have evaluated the effect of context-based process semantic annotation through a controlled experiment to test whether annotation can facilitate searching, navigating and understanding process models stored in a repository. For the evaluation we used the Method Evaluation Model (MEM), a widely accepted model for measuring the performance and user perception of artefacts. In order to perform the experiment we implemented a repository prototype that implements the annotation model and populated it with more than 100 process models.

The results provide evidence that the annotation model positively affects searching and navigating a process model repository. Furthermore, the results indicate that the annotation model positively affects the understandability of process models. However, the effect of the model on understandability is not significant. One of the reasons could be that most process models used for the experiment were not complex. Therefore, we hypothesize that CPSAM-based annotation could improve the understandability for very large and complex process models.

The results from the post-task survey suggest that most users perceived the annotation as easy to use and useful for searching, navigation and understanding of process models. Also, the results showed that users have positive intention to use the annotation model for navigation and understanding. However, most users are not sure about their intention to use the annotation model for searching. Possible reasons may include: a) annotation-based search requires some time and effort to think about different annotation elements before searching, and b) people are used to keyword-based search in their searching routine.

One of the limitations of the study is that a small number of participants were used for the experiment. Future research should aim at a large-scale evaluation of the annotation model and to improve the annotation model based on the evaluation results.

8 Business Process Relationship: The Meta-model

8.1 Overview

This chapter presents a process relationship meta-model as a template or blueprint for identifying and defining the relationship between process models in the repository. The purpose of the meta-model is to improve process model retrieval by enhancing the user's initiated navigation tasks where the user retrieves the complete set of relationships for a particular process model, inspects the relationships, selects the targets and retrieves the candidate process model. The meta-model is based on existing process relationships in the literature and the process-assets and asset-processes archetypes we have developed to help organizations find all business processes that exist in an enterprise. The chapter builds on Papers VI and VII as listed in section 1.5 of Chapter 1. The chapter begins by presenting the requirements that must be fulfilled by the meta-model in order to address the retrieval problem. We discuss how the relationship meta-model was developed following a three-phase process – identification of process relationship concepts, validation of the relationship concepts and construction of the meta-model. We then discuss the process-assets and asset-processes archetypes as the main concepts from which the meta-model is based. The chapter then present the study carried out to validate the archetype through a real world case study. The chapter describes the business process relationship meta-model with a description of each element and relationship types. The evaluation of the meta-model against established requirements is also discussed. The chapter ends with a summary and discussion about the meta-model.

8.2 Requirements for the Process Relationship Meta-model

The business process relationship meta-model serves as a common template or blueprint for identifying and defining the relationship between business processes in the repository. As a consequence, the meta-model has to fulfil a set of requirements. In accordance with the design science research process (Joahannesson & Perjons, 2012), we have identified a set of requirements for the process relationship meta-model. The requirements have been drawn

from the literature related to conceptual models (Krogstie and Sølvsberg, 2003), the conceptual model quality framework (Moody et al, 2003) and object repository system (Bernstein, 1997).

- *Relationship semantics* (Bernstein, 1997). Relationship semantics is the meaning of the relationship. A relationship without meaning is simply a connection that would be of no value. Process relationships are highly semantic elements, which express knowledge about business process interconnections. Therefore the relationship concepts that constitute the meta-model should express meaningful and useful process interconnections.
- *Comprehensibility* (Moody et al, 2003; Krogstie & Sølvsberg, 2003). The comprehensibility of the relationship meta-model is the ability to be comprehended by the users. The applicability and usefulness of the meta-model depend on the extent to which users are able to interpret and apply the relationship concepts in the repository. Therefore the semantics of the relationship concepts that constitute the meta-model should be understandable by non-modelling experts.
- *Domain independent* (Shahzad et al, 2010). One of the requirements for the proposed process repository is to store process models regardless of their domain. This means that the meta-model should have the right generality – general concepts used for identifying and defining relationships of business processes must be recognizable by users from any domain or discipline.
- *Completeness* (Moody et al, 2003). The meta-model is complete if it includes all the necessary relationships relevant for the repository use. There are no important relationships that are correct and relevant but missing from the meta-model. Therefore the meta-model should give a complete representation of the business process relationships that are relevant for the repository.
- *Tool support* (Krogstie & Sølvsberg, 2003). It should be possible to instantiate the meta-model using a tool to facilitate the identification and definition of process relationships in the repository. Tool support could be a way to achieving syntactic, semantic and pragmatic quality (Krogstie, 2012). Therefore the meta-model should lend support to the automated tool or assist in support for reasoning.

8.3 A Process Relationship Meta-model Development

The process relationship meta-model has been developed in a three-phase process: (1) identification of potential relationship concepts, (2) validation of the relationship concepts, and (3) construction of the meta-model. This is related to the design and develop artefact activity of the research process shown in Figure 13.

Phase 1: Identification of process relationship concepts

In this phase, process relationship concepts were identified from two sources. The first source is the existing literature on process model repositories and the second source is our own research project (Bider et al, 2012), which initially aimed at developing a procedure to find all processes in an enterprise.

Step 1: Analysis of existing literature on process model repositories

Here we performed a literature analysis by surveying some comparable research efforts, particularly as related to process model relationships in the repository. Malone et al (2003) are among those who can be noted for their pioneering of the concept of defining the relationship between business processes in the repository. They define two types of relationship between business processes: (1) whole-part relationship, and (2) generalization-specialization relationship. The whole-part relationship exists when one process is composed of one or more processes (called subprocesses). An example is “Manage order approval”, which includes “Handle rejected order”, and the latter is a part of the former, whereas the generalization-specialization relationship exists when one process (called a specialization) specializes another process (called a generalization). An example is “Manage returns”, which is a generalization of “Manage returns with prior approval”.

Other work defining the relationship between business processes in the repository includes the work of Kurniawan et al (2012). In their work they identify three types of relationship between process models: (1) whole-part relationship, (2) generalization-specialization relationship, and (3) inter-operation relationship. The first two types of relationship are the same as those defined by Malone et al. (2003). The inter-operation relationship exists between two processes when there is at least one message exchanged between them and there is no cumulative effect contradiction between tasks involved in exchanging the messages.

The scope of the above relationships is limited to directly related processes. They don't consider interdependencies between indirectly related processes in an organization. For example, if an academic institution wants to run a specific research project process, there could be several other processes that are related to it. To run such a process you need the workforce (researchers), which needs to be acquired and maintained. Therefore a process of hiring researchers is related to the research project process. Once the researchers have been hired, some processes are needed to make sure they are available for doing the work. These might include several processes such as manage researcher performance and develop and train researchers.

Given that the approach of using process model repositories to support reuse of process models is moving forward, significant effort in developing a formal way to identify and define the relationship between process models in the repository is critical. We believe a holistic approach to identifying and defining process model relationships in the repository is needed. In our

view, we find that the concepts that we have discussed above can best be tied together through an understanding of how business processes are interconnected in the overall process architecture of an enterprise. A process architecture is a schematic that shows how the business processes of an enterprise are grouped and interlinked (Frolov et al, 2010; Flores et al, 2012).

Step 2: Analysis of process-assets and asset-processes archetypes

In the second step, we analysed the process-assets and asset-processes archetypes (Bider et al, 2012) we had developed as a procedure to identify all business processes of an enterprise.

The archetypes were developed based on the enterprise model of Bider et al (2011), which represents an enterprise as consisting of three types of component: assets (e.g. people, infrastructure, equipment, etc.), sensors and business process instances.

In answering the question of “how the business processes are interconnected in an overall process architecture of an enterprise”, we hypothesized that “business processes and their relationships can be uncovered via a specially design recursive procedure”. The procedure starts with identifying the visible part of the “processes iceberg”, a so-called main process – a business process that produce value for which external customers are ready to pay. Typical examples of main processes are product manufacturing (e.g. computer manufacturing) and service delivery (e.g. teaching and learning process at a university), whereas examples of the external customers could be customers of a private enterprise, or a citizen paying for services provided to the public. Once the main process is identified, one proceeds by following up assets that are needed to run the main process. In this case, an asset is anything tangible or intangible such as materials, staff or other resources that are needed for successful running process instances of a certain type (Bider et al, 2012; Bider et al, 2011; Elias et al, 2014). Each assets type requires a set of supporting processes to have the corresponding assets in “working order” waiting to be deployed in the process instances of the main process. Typical examples of supporting processes are human resources (HR) processes (e.g. recruitment of employees) that ensure the enterprise has the right people engaged in its main processes.

On converting the above description into a formalized procedure that could be used in practice, two archetypes were introduced: (1) *Process-assets archetypes* (patterns), which help to find out what assets are needed for a particular process, and (2) *Assets-processes archetypes* (patterns), which help to find supporting processes that are needed to have each type of asset ready and available for deployment.

Since relationship concepts drawn from the literature have been well accepted and used, only process-assets and asset-processes archetypes were validated. Therefore, as the next step, the archetypes were validated by domain experts in a case study.

Phase 2: Validation of the archetypes

In the second phase, we validated the archetypes to test the appropriateness of the archetypes for identifying and defining the relationship between business processes (Elias et al, 2014). As proposed by (Poels et al, 2005), the semantic quality must be evaluated before proceeding to the application or implementation of the model or a method as an early quality indicator of the system that implements the model. Semantic quality expresses the degree of agreement between the information conveyed by the model (in this case the archetypes) and the domain that is modelled. The validation was done in a real-world case study from a higher education institution in order to validate its applicability (Elias et al, 2014). For the validation, the archetypes were applied to design the process architecture of major business processes, in the department, related to teaching and learning. The produced education process architecture was then evaluated by the business domain experts. More specifically, the validation was aimed at testing the appropriateness of the archetypes for identifying and defining the relationship between business processes.

Phase 3: Meta-model construction

In the third phase, the meta-model was constructed. To construct the meta-model we integrated the directly related process relationships as identified from the analysis of existing process model repositories and the indirectly related process relationships. The indirectly related process relationships are based on the process-assets and asset-processes archetypes (Bider et al, 2012). From the analysis of existing relationships we adapted the generalization-specialization and the whole-part (also referred to as part-include) relationships. Following the analysis of the process-assets and asset-processes archetypes, we identified two main types of relationship: the support process relationship and manage-managed relationship.

In Section 8.4 below, we describe the archetypes and show how they can be applied to unveil the dynamic process structure of an enterprise. In Section 8.5, we present the validation of the archetype before describing the meta-model in Section 8.6.

8.4 The Process-Assets and Asset-Processes Archetypes

In this section we describe the process-assets and asset-processes archetypes that form the bases of the proposed process relationship meta-model.

8.4.1 The Process-Assets Archetype for Main Processes

We consider an enterprise to be any organization whose operational activities are financed by external stakeholders. For example, an enterprise

could be a private company that gets money for its operational activities from the customers, or a public institution that gets money from the taxpayers. We also consider a main process to be any process that creates value to the enterprise's external stakeholders, which they are willing to pay for. Our definition of the term main process may not be the same as those of others (Scheer, 2000; Hammer & Stanton, 1999). For example, we consider as main processes neither sales and marketing processes, nor product development processes in a product manufacturing company. However, our definition of the main process does cover processes of producing and delivering products and services for external stakeholders, which is in line with other definitions of main processes (Hammer & Stanton, 1999; Scheer, 2000).

Main processes are the vehicles of generating money for operational activities. To get a constant cash flow, an enterprise must ensure that new business process instances (BPIs) of main processes are started with some frequency. To ensure that each started BPI can be successfully finished, the enterprise needs to have assets ready to be employed so that the new BPI gets enough of them when started. We consider that any main process requires the following six types of assets (Bider et al, 2012) (see also Figure 39 and Figure 40):

1. *Paying Stakeholders*. An actor who pays money to the organization for the service or product offered in a business transaction. Whereas an actor is an entity such as a person or an organization involved in the realization of a business transaction (Dunn et al, 2005). Examples of the paying stakeholders are, customers of a private enterprise, members of an interest organization, local or central government paying for services provided for the public.
2. *Business Process Templates* (BPTs). A description or definition of how the work is performed in an organization. It can exist in explicit or implicit form or both (Bider et al, 2012). The explicit form of the BPTs could be a written document, process diagram, or computerized system. Whereas the implicit form of the BPTs are in the head of people. Examples of the BPTs are: For a software development company that provides customer-built software, BPT includes a software methodology (project template) according to which their systems development is conducted; For a production process in a manufacturing company, BPT includes product design and design of a technological line to produce the product; For a service provider, BPT is a template for service delivery.
3. *Workforce*. A trained and qualified person who is employed by an organization to run the main business process. Examples of the workforce are workers at the conveyor belt, physicians in a hospital setting and researchers in academic institution.
4. *Partners*. An external actor that has a collaborative agreement or working relationship with the organization. Examples of the partner asset are suppliers of parts in a manufacturing process, a lab that carries

out medical tests on behalf of a hospital. Partners can be other enterprises or individuals, e.g. retired workers that can be hired in case there is a temporary lack of skilled workforce engaged in a particular process instance.

5. *Technical and Informational Infrastructure.* The physical facilities, equipment and information required by an organization to successfully run the main process. Examples of this type of asset are production lines, computers, communication lines, buildings, software systems etc.
6. *Organizational Infrastructure.* The organization and business structures, procedures and policies necessary to run a business process. Examples of this type of asset are the organization management, departments, teams and policies regulating areas of responsibility and behaviour.

The type of process (main) together with types of assets required for running it constitute a process-assets archetype for main processes. An arrow from the asset to the process shows the needs to have these types of asset in order to successfully run process instances of the given type. A label on an arrow shows the type of assets. Figure 40 is an example of such instantiation for teaching and learning processes.

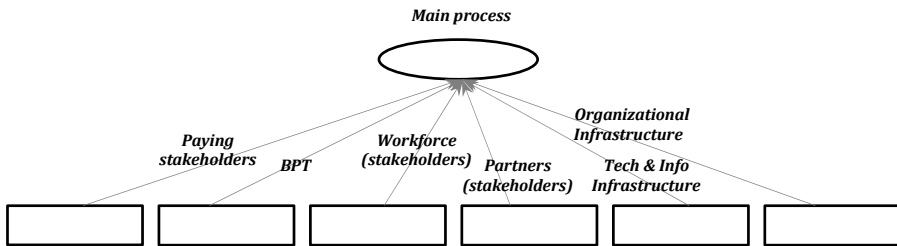


Figure 39: The process-assets archetype for main processes.

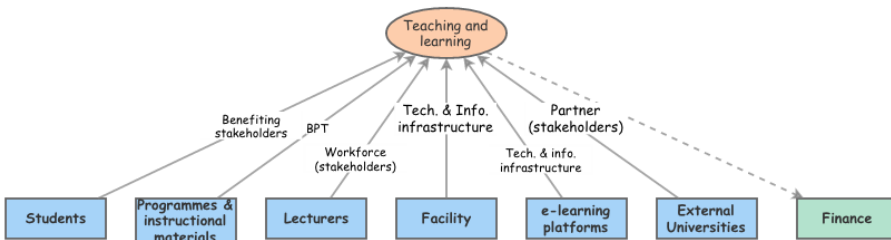


Figure 40: An instantiation of process-assets archetype for teaching and learning as the main process.

8.4.2 The Asset-Processes Archetype

In the previous section, we introduced six types of asset that are needed to ensure that business process instances (BPIs) of a main process run smoothly and with the required frequency. Each assets type requires a package of supporting processes to ensure that it is in a condition ready to be employed in BPIs of the main process. We present this package as consisting of three types of process connected to the life cycle of each individual asset (see also an example in Figure 41):

1. *Acquire*. Processes used by an enterprise to acquire or get a new asset of a given type. The essence of this process depends on the type of asset, the type of main process and the type of enterprise. For a product-oriented enterprise, *acquiring* new customers (paying stakeholders) is done through marketing and sales processes. *Acquiring* a skilled workforce is a task carried out inside a recruiting process. *Acquiring* a new BPT for a product-oriented enterprise is a task of new product and new technological process development. Creating a new BPT also results in introducing a new process in the enterprise.
2. *Maintain*. Processes used by an enterprise to keep existing assets in the right shape to be employable in the BPIs of a given type. For customers, it could be customer relationship management (CRM) processes. For the workforce, it could be training. For BPT, it could be product and process improvement. For technical infrastructure, it could be service.
3. *Retire*. Processes used by an enterprise to phase out assets that can no longer be used in the main process. For customers, it could be discontinuing serving a customer that is no longer profitable. For BPTs, it could be phasing out a product that no longer satisfies the customer needs. For the workforce, it could be termination or actual retirement.

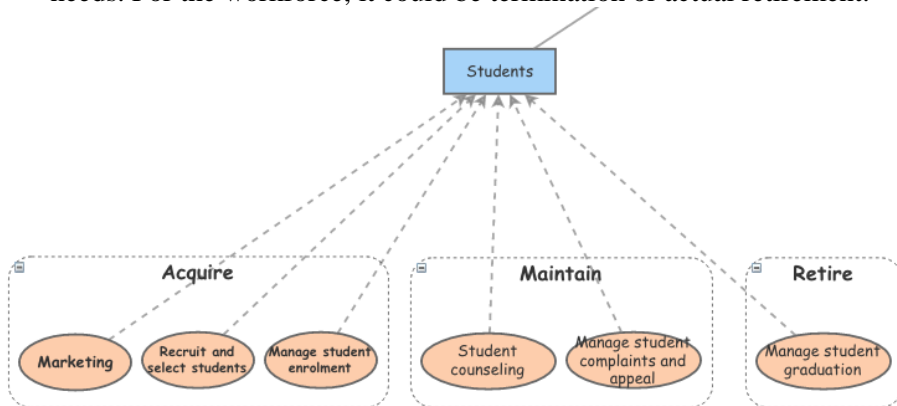


Figure 41: An instantiation of the asset-processes archetypes for "student" asset.

The asset-processes archetype can be graphically presented as in Figure 41. In the graph, the asset type is represented by a rectangle, and a process type by an oval. An arrow from the process to an asset shows that this process is

aimed at managing assets of the given type. The label on the arrow shows the type of process – *acquire*, *maintain* or *retire*. Instantiation of the archetype is done by inserting labels inside the rectangle and ovals. Actually, Figure 41 is an example of such instantiation for the *student* asset in a higher education institution.

8.4.3 Archetypes for Supporting Processes

Types of assets that are needed for a supporting process can be divided into two categories: general asset types and specific ones. General types are the same as for the main process, except that a supporting process does not need paying stakeholders. The other five types of assets needed for a main process – BPT, workforce, partners, technical and informational infrastructure and organizational infrastructure – might be needed for a supporting process as well.

Note also that some supporting processes, e.g. servicing a piece of infrastructure, can be totally outsourced to a partner. In this case, only the partner’s rectangle will be filled when instantiating the archetype for such a process.

Additionally to the five types of assets listed above, other types of assets can be added to a specific category of supporting processes. We have identified two additional assets for supporting processes of acquiring an asset that belongs to the category of stakeholders, e.g. paying stakeholders, workforce and partners:

Value proposition, for example description of products and/or services delivered to the customer, or salary and other benefits that an employee gets.

Reputation, for example of being a reliable vendor, or being a great place of work.

Adding the above two asset types to the five already discussed gives us a new process-assets archetype, i.e. the archetype for the acquiring stakeholders. An example of instantiation of such an archetype is presented in Figure 42.

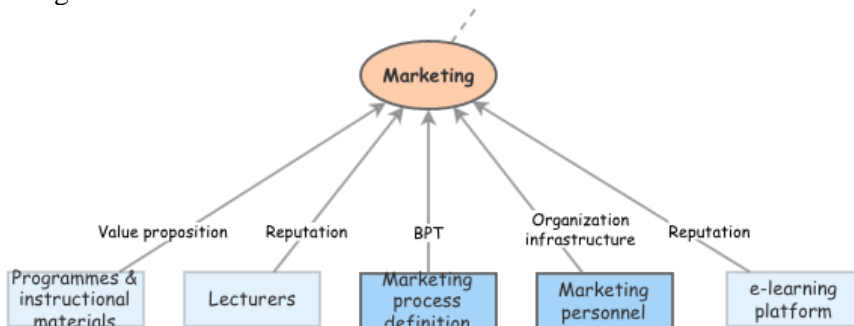


Figure 42: An instantiation of the process-assets archetype for acquiring students.

8.5 Validation of the Archetypes

In this section, we present an exploitation of the archetypes described in the previous section in a real-world case study from a higher education institution in order to validate its applicability (Elias et al, 2014).

For validation, the archetypes were applied to design a holistic view on the major business processes in the department related to teaching and learning and their interconnections. Then business domain experts evaluated the produced education process architecture that provides the holistic view of major processes. More specifically the validation was aimed at testing the appropriateness of the archetypes for identifying and defining the relationship between business processes. The following were investigated:

1. Whether the archetypes could be applied to build a process architecture in practice in a resource-efficient way
2. Whether the archetypes can reveal or explicate important facts about the business
3. Whether the archetypes could be understood and appreciated by domain specialists.

In the next subsections, we present the details of the study, which include: study settings, the validation process, the results and the analysis of the results achieved and lessons learnt.

8.5.1 Study Settings

In this section, we present the specific context in which the study was carried out.

The Organization. A case study to validate the archetypes was conducted at the Department of Computer and Systems Sciences (DSV). DSV belongs to the Faculty of Social Sciences at Stockholm University (SU) and carries out all types of academic activities – undergraduate, postgraduate and research with more than five thousand students. It runs several programmes in the fields of Computer Science and Information Systems. They include bachelor, master and doctoral programmes. DSV has more than two hundred staff members including teachers and administrative personnel.

The Process View. The core business of DSV is teaching and learning, research and consultancy. The focus of this study is on the teaching and learning as a primary business process performed in the department. Teaching and learning, as a main business process, involve all processes related to delivering knowledge to students. They include teaching, examining and graduation. We intended to investigate and map all business processes that are vital to the successful execution of the teaching and learning for knowledge delivery as the main business process.

The Team. The archetypes validation project involved senior staff from both teaching and administrative units; this group will be referred to as the *business domain experts*. The business domain experts included the director

of studies, director of finance and administration, head of academic units, IT director and coordinators of some specific academic programmes. The main team is that of *enterprise modelling experts*, which comprised of one PhD student and two teachers who had long industrial and academic experience in the field of enterprise modelling.

8.5.2 Archetypes Validation Process

The archetypes validation was a three steps process, as shown in Figure 43: (a) investigation of business processes, (b) designing the educational processes architecture of the department, (c) evaluation of the results.

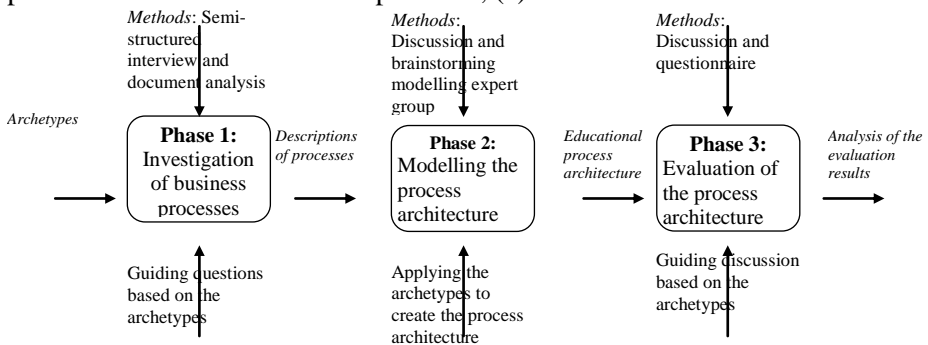


Figure 43: Archetypes validation process.

Step 1. Investigation of business processes

In the first step, as *modelling experts*, we investigated business processes performed in the department to enable the design of the educational process architecture. Here, we aimed at finding out what business processes are involved, and what is needed to successfully run the teaching and learning process at the department. Two methods were employed to investigate the business processes.

The first method was through semi-structured interviews (Denscombe, 2010). This enabled us to get an in-depth understanding of the business processes involved based on the issues brought up during the interview with the *business domain experts* (interviewees). We interviewed nine *business domain experts*, starting with the director of studies, who is the main senior person responsible for ensuring the successful execution of the teaching and learning process. The list of interviewees was later extended to the operational staff. Getting the input from the operational staff who perform the actual activities in different business processes increased our understanding of the details of various processes related to teaching and learning.

The second method used to investigate the business processes was document analysis (Denscombe, 2010). Document analysis focuses on information from formal documents or records. In this project document analysis was useful to complement the information obtained from the

interviews. Therefore during the interviews we received supporting documents, from which some of the business processes were identified.

The response of the interviews and documents were then analysed using a thematic analysis technique. The results of the analysis were used to design the educational process architecture described in the next section.

Step 2. Designing the process architecture

In the second step, we designed the educational process architecture by applying the process-assets and asset-processes archetypes and the data collected in the first step 1. It was during this step that the first goal of the archetypes validation project was achieved. The design was done by *modelling experts* using Insightmaker (Give Team, 2010) as a modelling tool. The choice was based on our knowledge of using the tool. The design process consisted of several iterations that can be presented as a sequence of the following sub-steps:

- Sub-step 1: In the first sub-step, we applied the process-assets archetype to identify the assets that are utilized by the department to ensure successful execution of the teaching and learning process. The resulting model is shown in Figure 40 and a detailed description of the model is provided in the results section, 8.5.3.
- Sub-step 2: In the second sub-step, we applied the asset-processes archetype to identify the processes involved for acquiring, maintaining and retiring each asset identified in sub-step 1. Figure 41 depicts the result of applying the asset-processes archetype for the asset student, the leftmost node of Figure 40. Similar results were produced by applying the asset-processes archetype to the remaining assets, i.e. lecturers, programmes and instructional materials, facilities, IT infrastructure (i.e. e-learning platforms) and external universities.
- Sub-step 3: Likewise, for each acquire, maintain and retire process identified in sub-step 2, we identified the assets needed for its execution. Again this was achieved by applying the process-assets archetype. Figure 42 is the results produced by applying the process-assets archetype to the marketing process, the leftmost node of Figure 41. Similar models were produced by applying the process-assets archetype to the remaining processes, i.e. recruit and select students, manage student enrolment, student counselling and manage student graduation.

Sub-step 2 and sub-step 3 are repeated for each asset and process respectively, until no more processes and assets could be identified.

Step 3. Evaluating the produced process architecture

In the third step, the educational process architecture, created in step 2, was used to evaluate the archetypes to determine their suitability. The evaluation was achieved through presentation and interview with seven business domain experts.

After completing the design of the process architecture, it was presented to the *business domain experts* (both teachers and administrative staff) individually. The presentation consisted of showing the process architecture. This was done first by showing how to start identifying the assets needed by the main process, in this case teaching and learning, and then going through the steps of identifying acquire, maintain and retire processes for each asset.

On completion of the presentation to individual business domain experts, the process architecture produced was validated through a semi-structured interview (Denscombe, 2010). The interviews were guided by an open-ended questionnaire. The results of validation are discussed in the next section, 8.5.3.

8.5.3 Educational Process Architecture: Results

The results of the archetypes validation project include the educational process architecture that shows all major business processes and their interconnection, and detailed description of each process performed at the department. Below we present and discuss a part of the educational process architecture.

8.5.3.1 Applying process-assets archetype for the main process

Figure 40 depicts the results of applying the process-assets archetype (sub-step 1 of Section 8.5.2) for the main process *teaching and learning*. From the study, it was learnt that several assets need to be available to successfully run the teaching and learning process at DSV. One of the primary assets that must be available for the main process to run is a student. The *student* is the benefiting stakeholder of the service being offered by the department. The main process is run by *lecturers*, a workforce asset that delivers knowledge to the student. The lecturers require *programmes and instructional materials*, an asset that includes the descriptions of programmes and courses, and instructional materials for each specific course.

Teaching and learning activities require infrastructure to make the process of learning effective. This infrastructure includes the teaching *facility* provided by the department, e.g. classrooms and offices, and the equipment required for teaching and learning. *IT Infrastructure* is another type of infrastructure utilized by the department to support teaching and learning. The IT infrastructure asset refers to the hardware, network resources, software applications, and services needed to support the teaching and learning process. The software applications include e-learning platforms.

As part of its organizational infrastructure, the department collaborates with other SU departments as well as *external universities* as partners towards achieving its goals.

In the next section we extend the results presented in Figure 40, by applying the asset-processes archetype to one of the assets, more exactly to the *student* asset.

8.5.3.2 Applying the asset-processes archetype for the student asset

Figure 41 depicts the results of applying the asset-processes archetype to the leftmost node of Figure 40. From the study, we have identified several business processes for acquiring, maintaining and retiring students.

Acquire Processes: Processes utilized by the department to acquire students include *marketing*, *recruit and select students*, and *manage students' enrolment*. The *marketing* process aims at attracting more students. It includes activities such as advertising education programmes in the Stockholm University (SU) catalogue and on the Web, and various marketing seminars. *Recruit and select students* aims at getting students for a specific academic year for various programmes offered by the department. It includes activities related to announcing, receiving applications, evaluating student applications, selecting students and sending offers to selected students. The *students' enrolment* process is related to administration of the accepted offers – once a student accepts the department's offer, the department has to manage their enrolment.

Maintain Processes: The processes for maintaining students include *student counselling*, which is designed to provide students with knowledge and awareness of options for selecting academic programmes or courses. Another process in this group is *manage student complaints and appeals*.

Retire Processes: *Manage student graduation* is the retire process, which includes activities related to application/petition for graduation, degree audit and course waivers and substitutions. It also includes activities related to notifying students of their graduation statuses, prepare graduation roster and certificates.

The material discussed in this section is part of the results produced by executing sub-step 2 from Section 8.5.2. Similar models were produced by applying the asset-processes archetypes to the remaining assets, i.e. lecturers, programmes and instructional materials, facilities, IT infrastructure (i.e. e-learning platforms) and external universities.

In the next section we extend the results presented in Figure 41 by applying the process-assets archetype to one of the processes, namely *marketing*.

8.5.3.3 Applying the process-assets archetype for the marketing process

Figure 42 depicts the results of applying the process-assets archetype to the leftmost node of Figure 41. From the study, we have identified that the marketing process at the department makes use of marketing materials, an asset which include information about *academic programmes offered* as a *value proposition*. The assets *lecturers* and *e-learning platforms* are used as the *reputation* of the department to attract students and stakeholders. To perform the marketing activities, the department requires *marketing personnel* and the *marketing process definition*, which describes the procedure of how the marketing process is to be performed. The assets *programmes and instructional materials*, *lecturers* and *e-learning platforms*

are general asset types because they are the same as for the main process (teaching and learning).

Figure 42 is the results produced by executing sub-step 3 of Section 8.5.2 to the marketing process. A similar process map can be produced by applying the process-assets archetype to the remaining processes, i.e. *recruit and select students, manage students' enrolment, student counselling and manage student graduation*.

8.5.4 Analysis of the Results

In this section, we present an analysis of the results and the lessons learnt from applying the archetypes. Specifically, the analysis aimed at establishing the following:

- (a) Whether the archetypes could be applied to build a process architecture in practice in a resource-efficient way
- (b) Whether the model can reveal or explicate important facts about the business
- (c) Whether the produced process architecture could be understood and appreciated by domain specialists.

The analysis is based on the following:

- Authors' own reflections on their experience from the project. This is used to answer the first question (a).
- Authors' own reflections, which were formulated as questions to business domain experts and then confirmed by the latter during the interviews after presentation of the process architecture to them. This answers the second question (b).
- Interview with business domain experts. This is used to answer the third question (c).

8.5.4.1 Efficiency of applying of the archetypes in practice

A detailed analysis of the results indicates that the archetypes can efficiently be applied to design the process architecture of an enterprise. We arrived at this conclusion by considering the following self-reflections.

Designing the educational process architecture consisted of two main steps: investigation of business processes and the actual design of the process architecture.

- *Investigation of business processes was the difficult part.* The investigation was done through interviews and document analysis. We interviewed nine business domain experts and the interview took approximately one hour for each participant on different working days, whereas the analysis of the interview took approximately 24 hours. The analysis of documents took approximately 20 hours. The investigation and modelling phases were iterative. During the design we also conducted some follow-up interviews for further clarification, which

took at least one hour. Therefore, interviews, analysis of the interviews and document analysis took approximately 80 hours in a span of two months.

- *Designing the educational process architecture.* Applying the archetypes to design educational process architecture was relatively easy. This took approximately 40 hours. The most important thing was to be able to identify the assets by applying the process-assets archetypes and apply the asset-processes archetype to identify processes for acquiring, maintaining and retiring the respective assets. Applying the modelling tool (Insightmaker (Give Team, 2010)) was also easy. In Insightmaker, a ghost concept is used to avoid criss-crossing of links and flows. Ghosting allows a reference to be made to a primitive in the model, which is shown with a partially transparent graphical style. The ghost primitive was quite convenient for identifying assets used in several different processes.

Most of the major business processes related to the teaching and learning process were well captured in the produced process architecture to a degree sufficient to provide a holistic view on business processes in the department. However, it was difficult to represent some of the processes and assets that could be of importance. One of the processes that were difficult to model is the process for acquiring the “alumni society”. Despite the fact that the department makes use of the alumni society as an asset for marketing, it was not clear how, by whom and when the asset is created and maintained at the department. This requires further investigations with business domain experts. Another process that was not captured is the process for acquiring financing. Despite the fact that the department makes use of the financial asset to acquire and maintain other assets such as lecturers, facilities and IT infrastructure, it was difficult to represent it in the process architecture. Finance as an asset is not directly needed to run the teaching and learning process, however it is produced since the department gets finances from the government by providing teaching and learning. Therefore, one possible way is to represent the finance as an asset produced by the teaching and learning process. With this, we propose using an arrow pointing from the teaching and learning process to the finance as an asset.

From these discussions, it is evident that applying the archetypes to design the process architecture of an enterprise does not require extensive resources. However, an initial effort is required to understand how a particular enterprise operates. Once one understands well the operation of an enterprise in question, the actual application of the archetypes to create the process architecture is relatively fast and easy.

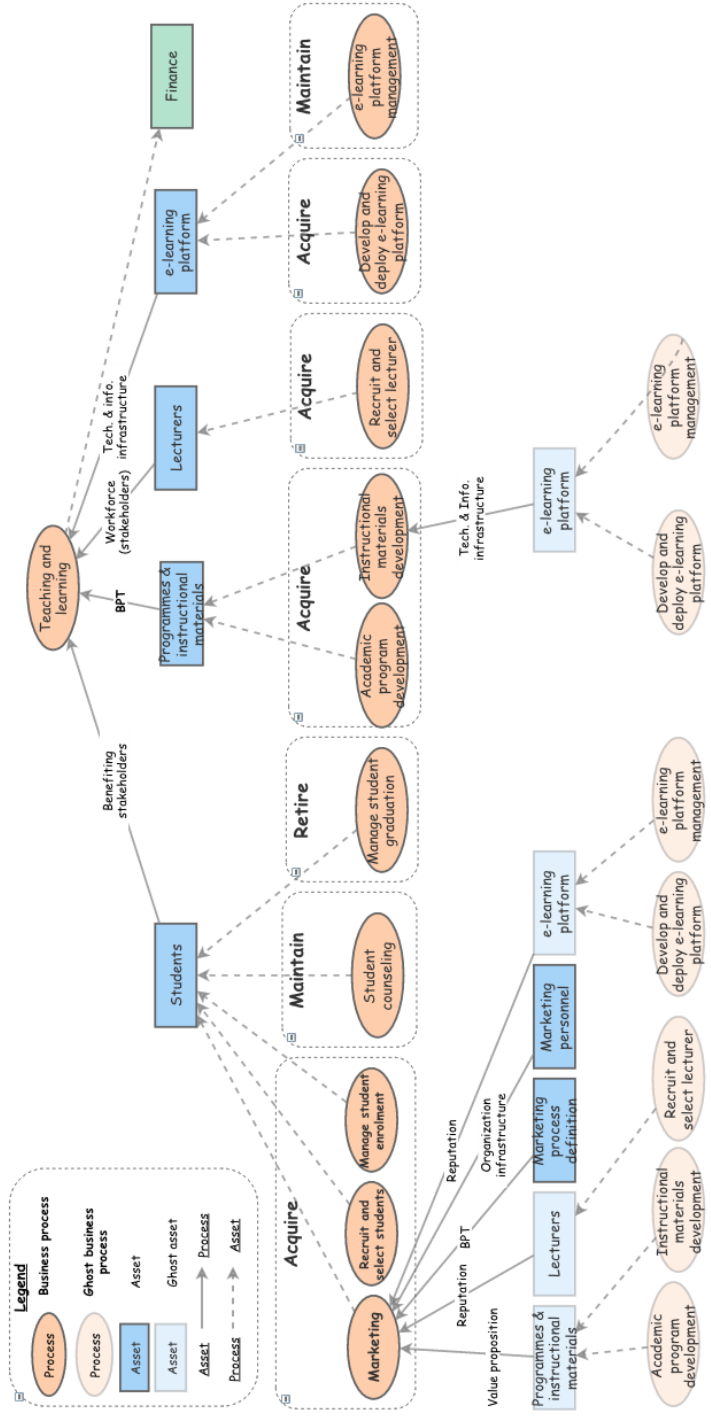


Figure 44: An instantiation of process-assets and asset-processes archetypes.

8.5.4.2 Important and useful business facts revealed by the archetypes

A detailed analysis of the results indicates that applying the archetypes to design the process architecture of an enterprise can reveal important and useful business facts. We arrived at this conclusion based on own reflections and validation of results by business domain experts through the interview.

The produced educational process architecture revealed several important and useful business facts. The following are some of the important facts revealed from the process architecture shown in Figure 44 (a part of the complete educational process architecture):

- Whereas the primary purpose of the lecturer recruitment process is to provide the department with qualified teachers, the process architecture reveals that the same process is intended to provide the marketing with the competitive advantage of having highly qualified staff, e.g. full professors, Nobel laureates, etc.
- Likewise, while the academic programme development is primarily aimed at providing the department with quality academic programmes and instruction materials for teaching and learning, the process architecture reveals that it also provides marketing with the competitive advantages of having attractive programmes and high-quality teaching materials.
- This was also the case with the develop and deploy e-learning platforms process. While the process aimed at providing teachers and students with a platform for educational delivery and management, the process architecture revealed that it also (1) provides teachers with tools for instructional materials development, and (2) provides marketing with the competitive advantage of having a high-quality educational delivery platform.

The above facts were validated by *business domain experts* through semi-structured interviews that followed the presentation of the process architecture. In addition, several other business facts were revealed during the interviews. For example, one of the experts revealed that the lecturer recruitment process is also intended to provide the department with highly qualified researchers for the research process. While the scope of the study only focused on the teaching and learning process, this revelation shows how the archetypes could be useful for providing a holistic view of all processes in an enterprise.

8.5.4.3 User comprehension and acceptance of the archetypes-based process architecture

A detailed analysis of the results indicates that process architecture produced, by applying the archetypes, is easy to comprehend and valued by

users. We arrived at this conclusion based on the responses from the interview with business domain experts.

The results of the interview with our *business domain experts*, which directly followed the presentation, show that the educational process architecture is well understood and appreciated by domain specialists. More specifically, the interviews were aimed at determining the following:

- Whether it is important to make explicit all purposes of all processes in the department. The results show that 100% (56% agree and 44% strongly agree) of the business domain experts agree that it is important to make explicit all purposes of all processes in the department.
- Whether the archetypes are useful for explicating business processes and their relationships. The results show that 100% agree that the archetypes are useful for explicating processes and their associated relationships.
- Whether the visual diagram that shows how all processes in the department are related is useful for business planning and development. The results show that 89% (56% agree, 33% strongly agree) agree that the visual diagram that shows how all processes in the department are interconnected is useful for business planning and development and 11% are not sure. When asked why he was not sure, an expert said he would be sure after applying the diagram to practical business development.
- Whether the presented visual diagram could be useful. The results show that 89% (56% agree, 33% strongly agree) agree that the presented diagram of showing business process relationships is useful for business planning and development and 11% are not sure. Similarly, when asked why he was not sure, an expert said he would be sure after applying the diagram in practice.

8.6 Business Process Relationship: Meta-model

In this section, we tie together all the concepts discussed thus far into a process relationship meta-model. Oeil2 et al (1992) define a meta-model as a “set of basic concepts which are related to each other, the so-called concept structure and a set of constraints determining the set of possible application models and the set of possible transitions between application models”. On the other hand, Lucena et al (2008) define a meta-model as a specification of elements of a language and the relationships among them. From these two definitions the following can be recognized about a meta-model. A meta-model provides: a set of concepts, a set of relationships and a set of constraints.

The process relationship meta-model is presented in Figure 45 below. The meta-model employs the concepts from process-assets and asset-processes archetypes described above and the process relationships defined in (Malone et al, 2003; Kurniawan et al, 2012). At the core of the

relationship meta-model is a business process. A business process may consist of other processes called subprocesses. A subprocess is a set of activities that have a logical sequence that meet a clear purpose. It is a process in itself, whose functionality is part of a larger business process. On the other hand, a business process may be specialized to another process called a “specialized process” that is an alteration of the former referred to as a “generalized process”.

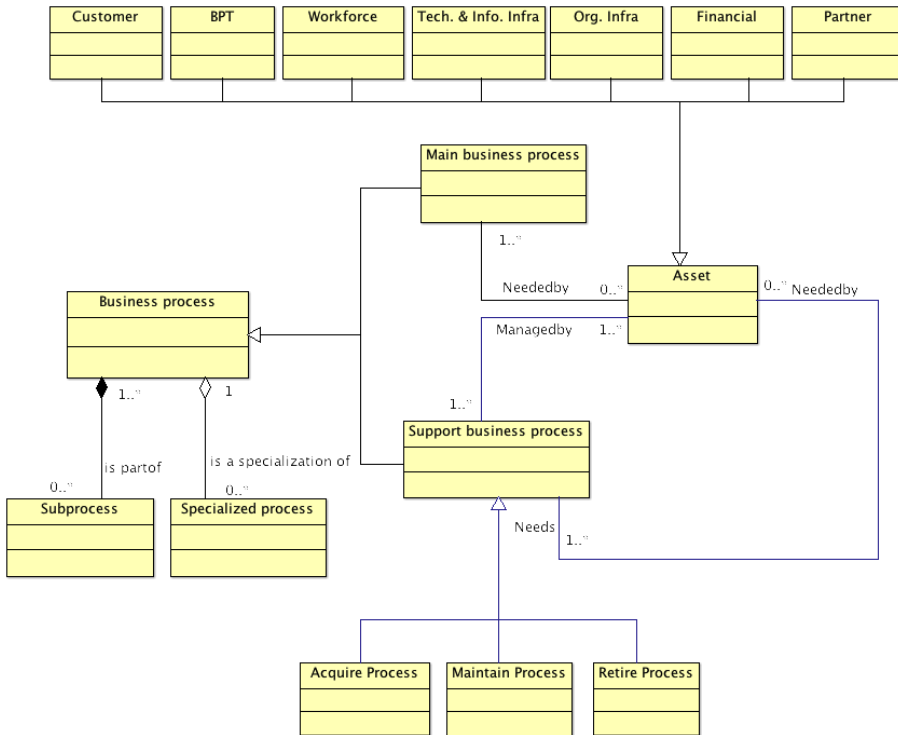


Figure 45: A business process relationship meta-model.

The business process is either a main process or a support process. A main business process is a process that produces value for the enterprise’s external stakeholders for which they are willing to pay, whereas a support business process is a process that produces value for the enterprise’s internal stakeholders. A *main business process* needs a set of *assets* for its successful execution. The assets needed by a business process are managed (acquired, maintained and retired) by *support business processes*. The support process may be of the type acquire, maintain or retire. The support process, like any other business process, may also need a set of assets for its successful execution. The assets needed by the support process cannot be the same as the assets it manages (acquire, maintain or retire).

In the repository, a business process may be related to a subprocess, or a specialized process. These are called “directly related processes”. On the other hand, a business process may be related to another process through an asset it requires for its successful execution. Therefore in the repository, every process will be associated with a list of assets it requires and for every asset defined in the repository a definition of the acquire, maintain and retire processes will be provided.

Below we present the relationships that may exist between business processes in the repository by referring to the meta-model.

Generalization-specialization relationship

A generalization-specialization relationship exists between two business processes if one process (called a “specialized process”) is an alteration of another process (called a “generalized process”). A specialized process is more specific and is fully consistent with the generalized process. An example is “recruit and select temporary lecturer”, a specialization of “recruit and select lecturer”, which is a generalization of the former. From our case study described in the next section, it was learnt that the department (DSV) could employ a lecturer on either a temporary or permanent basis. The normal (permanent lecturers) recruitment process is performed in cooperation with Stockholm University, where the recruitment of temporary lecturers is only carried out within the department.

Partof-include relationship

A partof-include relationship exists between two business processes if one process (called “whole process”) is composed of one or more processes (called subprocesses). A subprocess is a functionally complete and self-sufficient artefact for business process design and execution (Schumm et al, 2011a), whereas a whole process is more than the composition of the subprocesses. The subprocess has the partof role and the whole or parent process has the include or compose role (Elias et al, 2010). An example is “teaching and learning”, which includes “examination”, and the latter is a part of the former.

Support relationship

A support relationship exists between two business processes if one process (called a “support process”) manages an asset that is needed by another business process (called a “main process”) for its successful execution. The management of the asset by the support process may include acquiring, maintaining or retiring it. Therefore a support process can be an acquire or maintain or retire process type. An example is the relationship between “recruit and select lecturer” (as a support process) and teaching and learning process (main process).

A support relationship is transitive. If process P2 supports P1, and P1 supports P, then P2 supports P.

$$\forall p_2, p_1, p((\text{support}(p_2, p_1) \wedge \text{support}(p_1, p)) \rightarrow \text{support}(p_2, p))$$

An example in a higher education institution is the “develop and deploy e-learning platform process” supports the “instructional materials development process”, which then supports the “teaching and learning process”. Therefore the develop and deploy e-learning platform process supports the teaching and learning process.

A support relationship is shareable. Several business processes can be supported by the same process, i.e. two or more processes can share the same asset. An example is the “teaching and learning process” and “research project process” in a higher education institution. The two processes both need a lecturer as an asset and are therefore supported by the “recruit and select lecturer process”.

8.7 Evaluation of the Meta-model

The business process relationship meta-model has been developed for identifying and defining the relationship between process models in the repository. Such relationships facilitate navigation of the repository and enable users to locate related process models. In this section, the meta-model is evaluated against the requirements using an informed argument (Hevner et al, 2004; Peffers et al, 2012). Below we describe the evaluation.

Relationship semantics

The meta-model identifies four major types of process relationship that have specific meanings that are useful. The semantics of the part-include relationships also referred to as the whole-part relationship and the generalization-specialization relationship, have been well accepted. The new relationships we have introduced are those related to the archetypes: the support relationship and manage-managed relationship. The support relationship indicates that one process (called main process) requires certain assets, which are made available by a certain process, called the “support process”. The managed-manage relationship considers a business process as an asset, which needs to be managed by another process. From the case study, semantics of the interconnections provided by the produced educational process architecture were well accepted by the business domain experts of the department of DSV.

Comprehensibility

The meta-model consists of only eight concepts and four types of relationship, which are comprehensible to end-users. We identify two groups of users: those who create and define process relationships when populating the repository, and those who navigate the repository content based on the

defined relationships. Applying the meta-model to define process relationships is relatively simple as was learnt from the validation in the case study. However, the domain knowledge of the process in question is needed for the user who creates and defines the relationship during repository population or any other application of the meta-model. On the other hand, the lesson learnt from the case study, the process architecture produced, was well understood by non-modelling experts; this implies the concepts used to create the meta-model are comprehensible to all users. Therefore we argue that all the semantics of the concepts and relationships of the meta-model are straightforward to comprehend for both types of users.

Domain independency

The process relationship meta-model has been derived from the core components of a generic enterprise. The concepts business process, asset and asset type are common components in any enterprise. In addition, the meta-model incorporates generalization-specialization and whole-part relationships that are commonly used relationship concepts and semantics in the process repository as well as other application areas in different domains. While the domain knowledge of the business processes to be populated in question is important when applying the meta-model, the meta-model is not domain specific. It can be applied in any domain.

Completeness

The meta-model has extended existing process relationships by introducing two types of relationship based on commonly used concepts of a generic enterprise. There could be other types of relationship between business processes specific to particular domains, however our focus is on relationships that are meaningful, generic and domain independent.

Tool appropriateness

Our future goal is to implement the meta-model in the repository, however for the case study several tools were reviewed to define the educational process architecture. Most of the tools used in the case study, including the Insightmaker (Give Team, 2010), were found to be suitable for the work. Another suitable tool that was tested is Web Protégé – a free open-source collaborative ontology editor and knowledge acquisition tool for the Web (Tudorache et al, 2011). Therefore we argue that the meta-model can be instantiated by either applying an existing tool or by implementing it in the repository.

8.8 Summary and Discussion

In this chapter we have proposed a business process relationship meta-model, which can be used to identify and define relationships between business processes in the repository. This work builds on the requirements we introduced in Chapter 4 for process model repositories to support process model reuse and the review and analysis of existing repositories. One of the design issues is how to identify related business processes in the repository. To address this design issue, a business process relationship meta-model has been developed to help users to easily find related process models. The meta-model is intended to enhance the user's initiated navigation tasks where the user retrieves the complete set of relationships for a particular process model, inspects the relationships, selects the targets and retrieves the candidate process model.

The meta-model is made up of concepts taken from existing process relationships proposed for repositories and new concepts derived from process-assets and asset-processes archetypes we have developed to help identify all business processes in an enterprise. As a formal result of the proposed meta-model, the process relationship meta-model provides a common language or blueprint for identifying and defining relationships between business processes. The meta-model describes the relationship between process models as a way of representing all business processes in an enterprise and the interconnections between them.

The meta-model has been evaluated through an informed argument to test the extent to which it meets the established requirements. The results indicate that the meta-model meets well most of the established requirements and therefore it will enhance users' initiated navigation and enable related processes to be found in the repository. However, a practical evaluation of the model is required to test its performance in the repository.

9 Architecture for the Process Model Repository

9.1 Overview

This chapter presents the architecture of the process model repository that is publicly open and language independent, designed to support process model reuse. The main goal is not to suggest the best possible architecture, but to show that a good enough architecture can be designed based on the known architectural principles and knowledge sources. The principles and the knowledge sources were chosen based on their fitness for the task at hand, the main requirement being that they can be integrated in a reasonable whole that can be used for developing a repository. The design of the architecture is guided by the requirements discussed in Chapter 4 and the concepts discussed in the previous chapters. The chapter begins by discussing the design process of the repository architecture following a design science approach. We show how the requirements are mapped into technical problems, which are then decomposed into subproblems. The chapter presents a solution domain analysis carried out to identify architectural concepts to address each subproblem. We discuss how the identified concepts are then used to create the overall repository architecture by adopting service-oriented architecture. The specification of the repository architecture is presented based on ISO/IEC/IEEE 42010 standard architecture descriptions. We describe the architecture using the viewpoints according to the standard. We choose three viewpoints to describe the architecture: data and information, functional and standard viewpoints. The chapter also presents an architecture evaluation based on the analysis of the architecture against the requirements defined in Chapter 4. We discuss how each requirement is met by the architecture. A discussion of related work is also presented. The chapter concludes with some key remarks about the architecture.

9.2 Architecture Design and Development

In this section we describe the repository architecture design as the *design and develop artefact* activity of the research process shown in Figure 11. We choose Synthesis-based Software Architecture Design (SYNBAD) as an

approach for the design and development of the repository architecture (Tekinerdogan & Aksit, 2002). This is because in SYNBAD the scope of the architecture range from a systematic problem-solving perspective and not only from a stakeholder’s perspective. In the SYNBAD, the stakeholder’s perspective is abstracted to derive the technical problems. *Technical Problem* represents a specification of the problem to be solved. The process of deriving *solution abstractions* from *solution domain knowledge* is referred to as “solution domain analysis”. It relates to the concepts *solution domain knowledge* and *solution abstraction*, and the functions *search* and *extract* in Figure 46. The solution control part of the synthesis model represents the quantification, measurements, optimization and refinement of the selected solution abstraction.

In the following sections we explain the repository architecture design and development activity that consisted of four subactivities: technical problem analysis, solution domain analysis, create architecture and describe the architecture. The technical problem and solution domain analysis subactivities help us to identify architectural components and establish relationships between them. The create architecture subactivity applies the identified concepts to create the overall architecture of the repository.

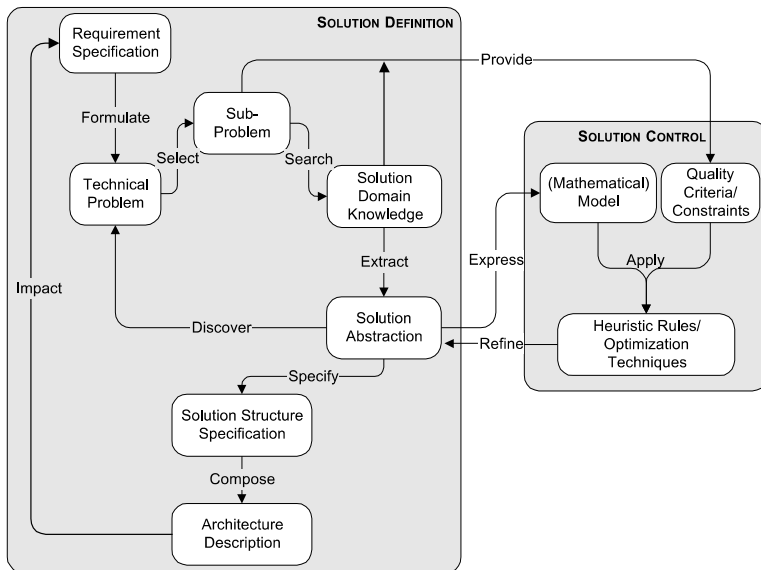


Figure 46: The Architecture Synthesis Model (Tekinerdogan & Aksit, 2002).

9.2.1 Technical Problem Analysis

In this subactivity, the requirements specifications (described in Chapter 4) are generalized and then mapped into technical problems. The technical problems are then decomposed into subproblems, which are also specified.

Before moving to solve the problems, however, prioritization is carried out to determine which subproblem needs to be solved first. The steps for this phase are described below.

Step 1. Generalizing and abstracting the requirements

System requirements are usually specific; as a result they provide only a specific interpretation of a more general problem. To provide a wider view and identify the problem, requirements specified in Chapter 4 were abstracted and generalized. The abstraction and generalization were done by using a thematic analysis technique (Ritchie & Spencer, 2002) to classify and identify the relationship between them. Requirements were categorized into six classes as shown in Table 16.

- *Process model representation*: The first category is process model presentation, which includes requirements (R1, R3, R5 and R7) aimed at providing users with better understanding of process models. While the primary purpose of annotating and classifying process models (R7) is to improve search and navigation of the repository, the annotation helps process model comprehension during the retrieval process. Therefore it is also included in this category.
- *Process model retrieval*: The second category is *process model retrieval*, which includes requirements (R7, R8 and R12) aimed at providing users with efficient retrieval mechanisms to enable them to find process models they want in the repository.
- *Process model management*: The third category is *process model management*, which describes the requirements (R6, R13, R14 and R15) aimed at providing the management of the repository content (process model) life cycle. These are fundamental requirements that are provided by any type of repository system.
- *Business process modelling*: Another category is *business process modelling*, which describes the requirement (R10) aimed at providing users with the mechanism to create new process models or edit existing process models within the repository environment.
- *Repository integration*: *Repository integration* is another category that describes requirement (R13) aimed at enabling easier integration of the proposed repository with external process modelling tools and external process model repositories.
- *Repository extensibility*: The final category is *repository extensibility*. This category describes requirement (R12) aimed at providing a repository that can easily be extended in terms of both *functionalities* and *information model* without requiring much effort.

Some of the requirements were not included in the technical problem analysis because they are not architectural design related. They include requirement R2 (the repository should allow process models to be stored

regardless of their domain) and R4 (the repository should store both business and process models).

Table 16: Generalized requirements (technical problems)

Requirements		Generalized requirements	Subproblems
R1	<i>The repository should support at least one standard process modelling notation.</i>	Process model representation	P1, P2
R3	<i>Process models in the repository should be represented in both graphical and textual form.</i>		
R5	<i>In the repository a business process should be represented by several process models with different levels of detail.</i>		
R7	<i>A process model should be annotated with information that can facilitate searching, navigating and interpreting process models.</i>		
R7	<i>A process model should be annotated with information that can facilitate searching, navigating and interpreting process models.</i>	Process model retrieval	P3, P4
R8	<i>Process models in the repository should be categorized based on different classification schemes to facilitate navigation.</i>		
R9	<i>The retrieval system should enable users to analyse and compare between process model similarities.</i>		
R10	<i>In the repository the relationship between process models should be defined and maintained to enable users to identify and find process models that are related to a candidate process model.</i>		
R6	<i>The repository should allow multiple variants of a process model to be maintained.</i>	Process model and repository management	P5, P6, and P7
R16	<i>The repository should allow multiple users to access and edit the same process model at the same time.</i>		
R13	<i>Therefore the repository should provide security to control access to the repository and operations on process models to enable sharing and exchange of process models between contributors.</i>		
R15	<i>The repository should provide tools and mechanisms to easily manage changes of the repository's functional components and contents structure.</i>		
R11	<i>The repository should allow process models to be created or modified, without requiring additional software installation on the client</i>	Business process modelling	Included in P1

	<i>side.</i>		
R12	<i>It should be possible to access and share process models with external repositories.</i>	Repository integration	P8
R14	<i>It should be possible to extend the repository information model or the repository capabilities.</i>	Repository extensibility	P9

Step 2. Identifying the technical problems from the generalized requirements.

In this step, we consider the generalized requirements to represent the technical problems that are to be addressed. They are: to provide a better representation of process models, to provide efficient process model retrieval, to provide a mechanism to manage the process model life cycle in the repository, to provide a process modelling mechanism, to provide easy integration and an interoperable repository, and to provide an extensible repository design.

Step 3. Specifying subproblems by decomposing technical problems into subproblems

In this step, we decompose the technical problems into subproblems. Some of the technical problems (i.e. repository integration and repository extensibility) were not decomposed and some technical problems were integrated into other subproblems: for example, business process modelling is integrated into subproblem P1. The third and fourth columns of Table 16 provide the technical problems and subproblems respectively. Below we describes each subproblem:

P1. Provide a standard and graphical process modelling environment with online support

- Goal: to improve user understandability by providing graphical visualization of stored process models
- Goal: to enable users to create new process models or modify existing ones over the Web without requiring them to install a modelling tool in their computers

P2. Provide a mechanism to manage process model complexity

- Goal: to enable users to generate an abstract process view, from a detailed and complex process model that focuses on relevant aspects of a business process

P3. Provide a mechanism to annotate and classify process collection

- Goal: to facilitate advanced searching of process models, navigating the process model repository, and understanding of process models to complement the standard visual notation

P4. Provide an advanced retrieval and process similarity check mechanism

- Goal: to enable users to easily find relevant process models

P5. Provide a version management mechanism

- Goal: to enable tracking of lineage of changes of a process model in the repository by defining and managing process model variants and their mutual relationships
- P6. Provide a repository configuration mechanism
- Goal: to enable advanced users to easily make necessary configuration changes in the repository
- P7. Provide a security mechanism
- Goal: to ensure that only authorized users are allowed to access the repository content
- P8. Provide repository integration and interoperability mechanisms
- Goal: to enable sharing and exchange of process models with external process model repositories and external modelling tools
- P9. Provide extensible repository structure and design
- Goal: to provide an architectural design that makes it possible to extend the information model and repository capabilities in the future without having to redesign the whole system

Step 4. Prioritizing the subproblems

In this research project, our prioritization of the subproblems was based on the same order in which the requirements and technical problems were identified.

9.2.2 Solution Domain Analysis

In this subactivity, the model for extracting the architecture design solution is provided. For every subproblem defined in the technical problem analysis phase, we identify and prioritize the solution domains. Then the knowledge sources for each solution domain are identified, studied and prioritized. After studying and analysing the solution domain knowledge sources, the fundamental concepts are extracted from them. The concepts are then structured using relations that are derived from the solution domains. The activities continue with refining the concepts. The steps for this phase are described below.

Table 17: The solution domain for the subproblems

SubProblem	Solution Domain (SD)	
<i>P1</i>	<i>Web-based process modelling</i>	<i>SD1</i>
<i>P2</i>	<i>Business process model abstraction</i>	<i>SD2</i>
<i>P3</i>	<i>Semantic process model annotation</i>	<i>SD3</i>
<i>P4</i>	<i>Searching and navigation</i>	<i>SD4</i>
<i>P5</i>	<i>Version management</i>	<i>SD5</i>
<i>P6</i>	<i>Configuration management</i>	<i>SD6</i>
<i>P7</i>	<i>Access control</i>	<i>SD7</i>
<i>P8</i>	<i>Heterogeneous data exchange</i>	<i>SD8</i>
<i>P9</i>	<i>Extensible system design</i>	<i>SD9</i>

Step 1: Identification and prioritization of the solution domain knowledge for each subproblem

In this step, for each subproblem we searched for the solution domain that provides the solution abstractions. The search process included searching through the Web using Google Search and Google Scholar. The searching of the solution domains also included discussion with supervisors and the research project team. Table 17 above provides a list of the solution domains we have identified for each subproblem.

Step 2: Identification and prioritization of the knowledge sources for each solution domain

In this step, for each of the solution domains, identified in step 1, we identified and prioritized the knowledge sources according to their objectivity and relevancy. The identified knowledge sources consisted of project team members, literature and documentation on existing process model repositories. The knowledge sources were chosen based on their fitness for the task at hand, the main requirement being that they can be integrated in a reasonable whole that can be used for developing a repository.

Table 18: A selected set of the identified knowledge sources of the solution domain

Solution Domain	Knowledge Source (KS)	
SD1	<i>Oryx – An Open Modelling Platform for the BPM Community (Decker et al, 2008)</i>	KS1
SD2	<i>Business process model abstraction: definition, catalogue and survey (Smirnov et al, 2012)</i>	KS2
SD3	<i>Toward reuse of object-oriented software design models (Ali & Du, 2004)</i>	KS3
SD4	<i>MIT Process Handbook (MIT process handbook project, 2001)</i>	KS4
SD5	<i>Repositories and Object-Oriented Databases (Bernstein, 1997)</i>	KS5
SD6	<i>Repositories and Object-Oriented Databases (Bernstein, 1997)</i>	KS5
SD7	<i>Repositories and Object-Oriented Databases (Bernstein, 1997)</i>	KS5
SD8	<i>Integrating Learning Object Repositories Using a Mediator Architecture (Kärger et al. , 2006)</i>	KS6
SD9	<i>Service-Oriented Architecture (SOA): Concepts, Technology and Design (Erl, 2005)</i>	KS7

Table 18 provides a selected set of knowledge sources for the overall solution domain an “open, integrated and language-independent process model repository”.

The knowledge source for extensible system design was more difficult to find, however the literature indicates that adopting Service-Oriented Architecture as the architectural style enables extensibility and scalability (Erl, 2005). The knowledge source for solution domains SD5, SD6 and SD7 is the same.

While in Table 18 we have shown only one knowledge source for each solution domain, several knowledge sources were identified from which one knowledge source was selected for the overall solution domain; for example, Table 19 shows a selected set of knowledge sources for the solution domain (SD2) *business process model abstraction technique*.

Table 19: A set of knowledge sources for solution domain business process model abstraction technique (SD2)

Solution Domain	Knowledge Source	Form
SD2	<i>Business process model abstraction: definition, catalogue and survey (Smirnov et al, 2012)</i>	<i>Journal paper</i>
SD2	<i>Flexab – Flexible Business Process Model Abstraction (Weidlich et al, 2011)</i>	<i>Conference paper</i>
SD2	<i>Business process model abstraction (Polyvyanyy et al, 2010)</i>	<i>Book chapter</i>
SD2	<i>Business process model abstraction: Theory and practice (Smirnov, 2010)</i>	<i>Book</i>
SD2	<i>A semantic approach for business process model abstraction (Smirnov et al, 2011)</i>	<i>Conference paper</i>
SD2	<i>The Triconnected abstraction of process models (Polyvyanyy et al, 2009)</i>	<i>Book section</i>

Step 3: Extract solution domain concepts from solution domain knowledge

In this step, the identified knowledge sources were studied and analysed to extract the fundamental concepts. After deliberating on commonalities and variability of the extracted information from the knowledge sources we could extract the following concepts:

- *Process model editor and process modelling engine*. KS1 is the knowledge source for the *Web-based process modelling as a solution domain (SD1)*. KS1 describes a business process modelling platform that enables users to create process models on the Web. From this knowledge source we can extract two solution domain concepts: *process model editor* and *process modelling engine*. The *process modelling engine* is a system component that implements the mechanism that provides process the modelling environment in the repository.
- *Process model abstraction engine*. KS2 is the knowledge source for the *business process model abstraction as a solution domain (SD2)*. KS2 describes process abstraction techniques and how business process models can be abstracted from detailed complex models to less complex process models based on different use cases. The core component of the

process abstraction techniques is the abstraction engine. Therefore from this knowledge source we can extract a *process model abstraction engine* as solution domain concepts. It is a system component that implements the mechanism for abstracting complex process models to less complex process models.

- *Semantic annotation manager*. KS3 is the knowledge source for the *semantic process model annotation* as a solution domain (SD3). This knowledge source is based on the researchers' knowledge and experience and review of how the annotation of software components is applied to improve retrieval of reusable software objects. From this knowledge source we propose a *semantic annotation manager* as a solution domain concept. It is a subsystem for semantically annotating process models to facilitate efficient retrieval of process models.
- *Process information mediator*. KS6 is the knowledge source for *heterogeneous data exchange* as a solution domain (SD8). It describes how a repository of learning objects can be integrated with various external repositories to enable access and sharing of learning objects through the use of mediator components. From this knowledge source we extract *process information mediator*, which serves as an interoperability layer to enable integration and exchange of process models between the proposed process model repository and external process model repositories.
- *Repository engine*. KS4 is the knowledge source for the solution domain searching and navigation (SD4); and KS5 is the knowledge source for the solution domain version management (SD5), configuration management (SD6) and access control (SD7). KS4 describes the process model repository, which provides the searching and navigation mechanisms. KS5 describes the design and functionalities offered by the object repositories. From this knowledge source we extract a *repository engine* as the solution domain concept, which provides a mechanism for overall repository management including *advanced searching and navigation components*, *version manager*, *access control manager* and *configuration manager*.
- *Web services and Enterprise Service Bus (ESB)*. KS6 is the knowledge source for *extensible design* as the solution domain (SD9). It describes how to design an extensible system by following a service-oriented approach as architectural style. From this knowledge source we extract *SOA-based concepts such as Web services* and enterprise service bus (ESB) as solution domain concepts.

Step 4: Define conceptual structure

In this step, we make use of different structural relations to structure the identified solution domain concepts. Figure 47 depicts a conceptual structure

of the repository system flow chart. The conceptual structure of the system outlines clearly the interactions between system components along with the user activities. The rectangles depict functional components of the repository and the round-edged boxes describe the actions taken by the user. The dotted lines correspond to user actions, while the solid ones correspond to repository system flows. Populating the repository is accessible only to authorized users. Users need to be registered to exploit the search and retrieval features of the repository along with a set of advanced functionalities, including similarity check and model abstraction. Unauthorized users can only perform simple search requests to find the desired process model.

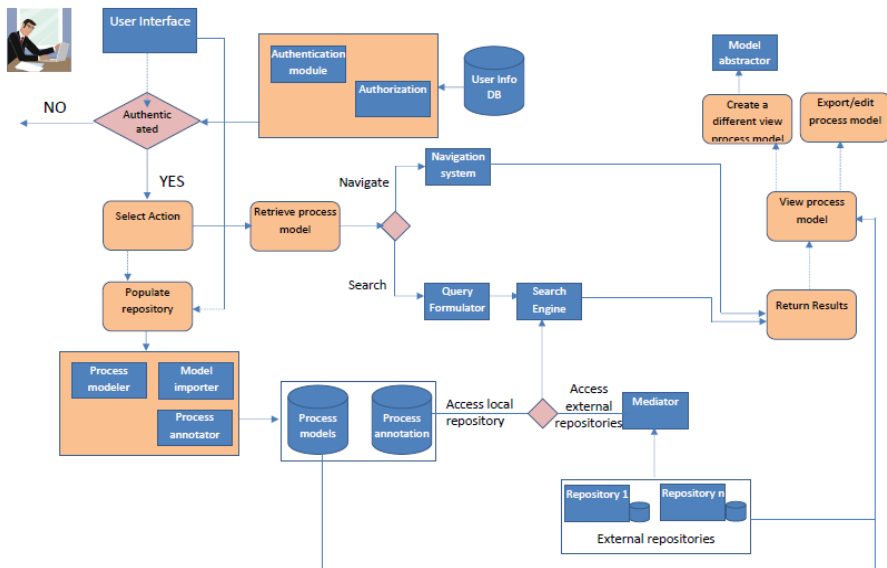


Figure 47: Repository system flow chart.

9.2.3 Creating the Repository Architecture

The technical problem analysis and solution domain analysis formed the basis of the overall repository architecture by identifying the architectural components and the relationships that exist among them as shown in Figure 47. In this section the identified concepts are used to create the architecture of the proposed process model repository.

Based on the understanding of the requirements and the relevant system specifications, it is obvious that the proposed repository consists of subsystems and heterogeneous components (i.e. existing repositories), relying on different technologies and implementation techniques. Furthermore, one of the major objectives of the proposed repository design project is flexibility and scalability, so that future modifications can easily add value.

The above observations signify the need to adopt a distributed architectural approach for the system. In this case, the best architecture style would be SOA (Service-Oriented Architecture).

9.2.4 The Repository Architecture

The architecture of the proposed repository shown in Figure 48 is composed of four layers: data, core, interoperability and presentation. The data layer corresponds to the process knowledge base, which stores all the process models, process annotation, process patterns, annotation model schema and user data. It constitutes the data access and data source components. The core layer corresponds to the process knowledge creation and reuse. It

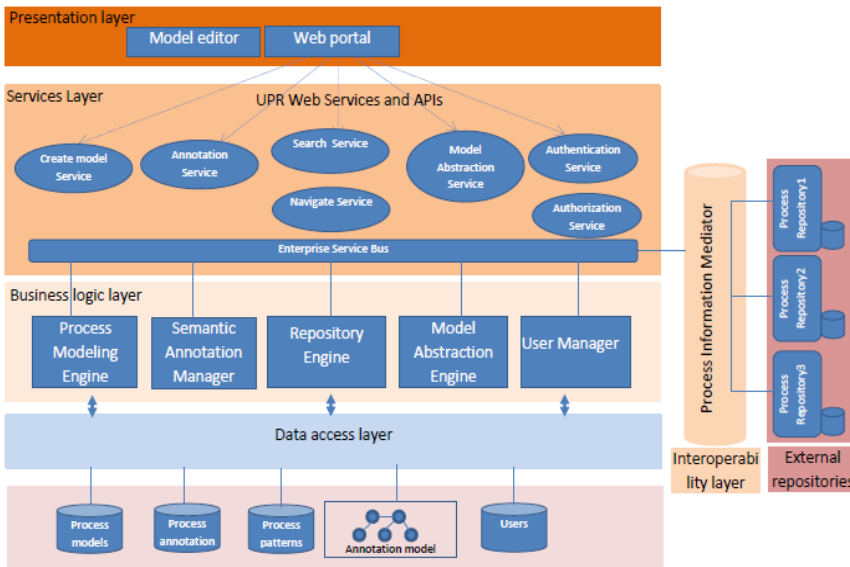


Figure 48: Architecture of the process model repository.

constitutes repository foundation services and repository service components, which implement the capabilities necessary for creating, manipulating and reusing the process knowledge. The repository service components include components for a repository engine, process modelling engine, semantic annotation manager and business process abstraction engine. The interoperability layer encompasses the services that implement the mechanisms and an interface for external integration and accessing multiple heterogeneous data sources of the process model repositories. The presentation layer encompasses the applications that use the functionalities provided by the repository through the Web service. The Web service provides an API that maps to the repository API and interfaces all the functionalities of the repository. The use of the Web service as an interface to the repository enables extensibility and integration. All the

communications between services and service components in the business logic are carried out through the enterprise service bus.

In the following section, we describe the architecture design shown in Figure 48 by using data and information, functional and standard viewpoints (ISO/IEC/(IEEE), 2011).

9.3 Architecture Specification

In this section the architecture is described as a subactivity of the design and develop artefact activity of design science. The architecture is described based on ISO/IEC/IEEE 42010 *systems and software engineering – architecture description* (ISO/IEC/(IEEE), 2011).

Based on the repository requirements and architectural concept identified in the previous section, three major viewpoints and their associated concerns are described below. They are the data and information, functional and standard viewpoints. In addition, UML activity diagrams have been applied to describe the main workflows of the process model repository system.

9.3.1 Data and Information Viewpoint

9.3.1.1 Conceptual Model

Process models: The main data entity stored in the proposed process model repository is the *business process model*. Mainly process models will be described using a standard modelling notation. However, in order to achieve a mass collection of process models, the repository will accept process models in any format that a contributor can provide. Therefore process models can be represented in different formats such as text documents and bitmap pictures of process models.

Process annotation: In order to ensure effective retrieval for a very large collection of process models with heterogeneity, process models will be annotated with semantic information. The annotation information will facilitate searching, navigating and understanding of process models. Therefore, besides storing process models the repository stores *process annotation*. The process annotations are based on the context-based process semantic annotation model (CPSAM) (Elias & Johannesson, 2013) we have developed. The process annotations are the instances of the CPSAM associated with stored process models. A detailed discussion of the semantic annotation model is given in Chapter 6 of this thesis.

Process patterns: In addition, the repository stores process patterns. These patterns are used to build up process models. Process models can either be of

a more general character like reference models or they can be specific solutions.

9.3.1.2 Data Access Model

All access to repository data is implemented as capabilities via a data access layer. Data access capabilities determine data access rights, based on user access scenarios in the repository. There are three classes of data access capability: create (adding process model), query and update. The data access layer manages exchanges with other systems and handles the transformation of relational data to XML where needed.

9.3.1.3 Data Storage Model

In the repository, process models can be stored as files or relationally, whereas the process annotations are all stored relationally. Applications have no direct access to data. All access is through the data access layer. The layer takes care of storing data from the upper layers to the database and retrieves it back, alongside updating and deleting it.

9.3.2 Functional Viewpoint

The proposed architecture consists of several components organized into layers. Below we identify and describe the functionalities of major components of the architecture.

9.3.2.1 Repository Engine

The repository engine is a repository component that provides fundamental functions for storing and retrieving process models and maintaining the relationships among them. In order to provide the storing and retrieval functionalities the repository engine implements the following services:

Search Service. A major goal of reuse is, of course, to “find the artefacts faster than the time it takes to develop them” (Krueger, 1992). Therefore, a mechanism to enable faster retrieval of the relevant artefacts is one of the main requirements for any reuse-based repository. The search module implements the search mechanism and provides the interface for enabling users to search and retrieve process models stored in the repository. The proposed repository provides two possible ways to search for a process model: (i) *keyword-based search*: a traditional searching method that allows a user to submit keywords and the repository retrieves related process models; (ii) *annotation-based search*: people who model processes as well as those who formulate the search queries might use different keywords to express the same concepts. Therefore it is difficult for the user to find the relevant process model by using keywords. In order to enable retrieval of relevant process models, the repository implements annotation-based search, based on the semantic annotation model (CPSAM) (Elias & Johannesson, 2013). It allows users to direct their search by using elements of the

annotation model. The repository retrieves all the process models that have been annotated with instances of one or more of the selected elements.

Navigation Service. Navigation service implements the navigation mechanism and provides the interface for enabling users to find a relevant process model easily through navigating the repository. A crucial challenge in designing navigation mechanisms is to create a navigation structure. A navigation structure determines the possible sequences for accessing process models, and imposes an organized layout on the repository's content (Huang, 2003). The navigation structure of the proposed repository is implemented based on the annotation model (CPSAM) (Elias & Johannesson, 2013). Therefore, the process models stored in the repository are organized by elements of the annotation model. The main advantage of navigation is that it allows users to evaluate a large number of process models rapidly and determine which is useful (Huang, 2003). Therefore, navigation is particularly useful when users don't know exactly what they need.

Versioning Service. The versioning service implements the mechanism for managing process model variants and the change history of the process models. Versioning is provided when a new process model is stored or an existing process model is modified. To record the change history, created or modified process model is stored as a new version in the repository. The version management method adopted in this repository is based on semantic annotation.

9.3.2.2 Process Modelling Environment

The proposed repository aims at providing a place where potential users can share process knowledge for reuse. Therefore an environment where users can create new or modify stored process models is necessary. In the repository, the process modelling environment implements and provides a process modelling editor through which process models can be created or modified with standard modelling notation. In addition, it enables graphical representation of business processes stored in the repository. The modelling environment consists of two major components: the process model editor and the process modelling engine (PME).

Process Model Editor: The process model editor, at the presentation layer, is a Web-based model editor that enables users to author processes graphically.

Process Modelling Engine: The process modelling engine (PME), at the core layer, is the core component that implements a process modelling subsystem. It constitutes a model converter and modelling language definitions (stencil sets). At the presentation layer business processes are represented in graphical form (called the "process model") using standard modelling notation. However, at the data layer process models are stored as process description (i.e. XML format). Therefore a model converter is responsible for converting process models into process description in XML

format and vice versa. The process models in XML format can be stored in a relational database or in a file system, such that they are easily accessible and imported into any Java IDE.

The modelling language definition (stencil sets) is another component of the PME. A process modeller can be built to support any selected standard modelling notation such as BPMN, EPC, UML AD, etc. Therefore a process modelling engine maintains process modelling definitions (stencil sets) for a specific modelling notation. Multiple modelling notations can be supported by creating and extending their definitions (stencil sets) (Decker et al, 2008) in the process modelling engine.

9.3.2.3 Semantic Annotation Manager (SAM)

In order to facilitate searching, navigation and model comprehension, the repository requires process models to be semantically annotated. Therefore, a mechanism for annotating, indexing and storing process annotation is required. The semantic annotation engine implements the mechanism to annotate, submit and index new process models in the repository (Elias & Johannesson, 2013).

There are two phases of process annotation: automatic annotation and manual annotation.

When a user completes modelling a business process, the create model service (save function) will store the model (in the process models database) and invokes the annotation service. The annotation service, during the automatic annotation phase, creates and indexes the basic annotations, which include process name, process description and the address of the stored model and stores them in the process annotation database. The annotate service then, during the manual annotation phase, lets a user annotate the process based on the CPSAM annotation model. The specific process annotation is then indexed and stored in the process annotation. The annotation service is also invoked whenever a process model is imported.

The CPSAM suggests a number of high-level instances or categories for each element, i.e. for resource we have goods, services, etc. These categories can be specialized and complemented with additional categories depending on the domain under consideration. Therefore, in order to support extensibility, the architecture includes an editor as part of the annotation service that allows users to define additional instances or categories for the included elements. The editor allows changes to be made only to the CPSAM schema in the data layer. The repository enables automatic updating of all repository services that are based on the annotation model. In this way, the extension does not break existing repository services that are oblivious to the extension, and also services that are aware of the extension are able to exploit it.

9.3.2.4 Model Abstraction Engine (MAE)

When modelling business processes, for each modelling goal, a specific process model at a certain level of granularity can be created. Therefore stored models may be complex for some use cases. Thus, a mechanism to create and provide different process views from detailed process models is necessary. The process model abstraction engine (MAE) provides and implements a mechanism for abstracting detailed process models into more coarse-grained models for specific use cases.

The MAE component consists of the abstraction service and the use case catalogue. The abstractor defines abstraction algorithms for different use cases, whereas the use case catalogue predefines use cases for which the model is intended. For example, a user may need fast process comprehension, a process model with relevant activities, or relevant process instances, etc.

During the model retrieval process, the process model editor of the PME presents the model to the user. If a user decides to generate different process views that meet a desired need, the abstraction service is triggered that takes the desired use case definition from the use case catalogue sent with the initial process model to the abstractor. Given this input, the abstractor performs the abstraction algorithm and produces an abstract model. The resulting model can then be stored as a version of the initial process for future reuse.

9.3.2.5 Process Information Mediator (PIM)

Our approach aims to enable access of process models from multiple repositories. However, each repository uses a different metadata structure to annotate process models and different technologies for storing them. This makes it almost impractical to retrieve and reuse process models from external repositories. The architecture introduces a process information mediator (an interoperability layer) that implements the mechanisms and a uniform interface for accessing multiple heterogeneous data sources. A query-rewriting mechanism based on metadata semantics enables the access of process models from external repositories (Kärger et al, 2006; Wiederhold, 1992). The process information mediator consists of several components, as shown in Figure 49.

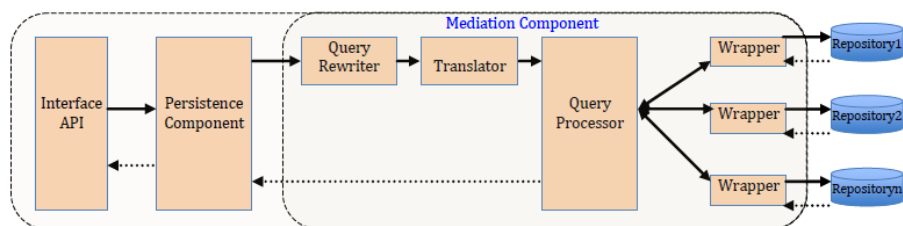


Figure 49: Process information mediator.

Interface Component: all interactions with the mediator are received via the interface API. It exposes an API that allows other services to make service requests through the mediator. It defines and manages all incoming query requests. The query requests are then passed to the persistent component for processing.

Persistence Component: the persistence component receives the query request from the interface API and starts and manages the execution of that request. It maintains the copy of each query received by the mediator and maintains the persistence data store for the request data, response data and the metadata for each query.

Query Processor: To allow users to query several process model repositories, the mediator implements the query processor, which accepts queries formulated in a uniform query language and returns the URIs, which points to a process model that meets the query. The query processor performs two services: query distribution and results integration. It receives the user query (in the form of the CPSAM) and distributes the query to each wrapper for translation. It then receives and integrates the query results from the wrappers and forwards them to the service layer for the user.

Wrapper: a wrapper for each connected repository is provided. The wrapper is responsible for transforming the query from the query processor into a format understandable to the target repository. The query sent to the wrapper will contain terms from the CPSAM (Elias & Johannesson, 2013). The terms of the CPSAM in the queries have to be replaced with the corresponding terms that a target repository uses to describe its process models. In order to ensure correct substitutions, a mapping between the terms of the metadata structure of the proposed repository and the terms of the metadata structure of the corresponding repository is defined and maintained in each wrapper. This is how the semantic interoperability is achieved.

9.3.3 Standard Viewpoint

The proposed process model repository may involve many entities. The need to integrate with external and existing repositories requires adherence to standard architecture to facilitate interoperability. There are several levels in which interoperability needs to be provided by the repository. Therefore to achieve these interoperability requirements the repository includes the following:

- *Terminology services:* The terminology services define a common terminology.
- *Web services:* The repository achieves interoperability between interfaces of different system components of the repository through the use of Web service standards.

- *Wrappers*: The repository achieve interoperability between the proposed repository and the heterogeneous data sources from external repositories through the use of wrappers, which translate the metadata semantics of the proposed repository to the semantics used by respective repositories.

9.3.4 Workflow Activities

Below we make use of the UML activity diagram to describe some essential activities of the system from the end-user perspective. They are “populating the repository” and “search and retrieval of process models”. The search and retrieval is described from two perspectives: accessing internal repository and external repositories.

9.3.4.1 Populating the Repository

There are three possible ways in which the proposed repository could be populated: (1) by importing process models; (2) by creating process models using the process modelling tool provided by the repository process modelling engine (PME); (3) also the repository could be populated automatically whenever access is made to the external repositories.

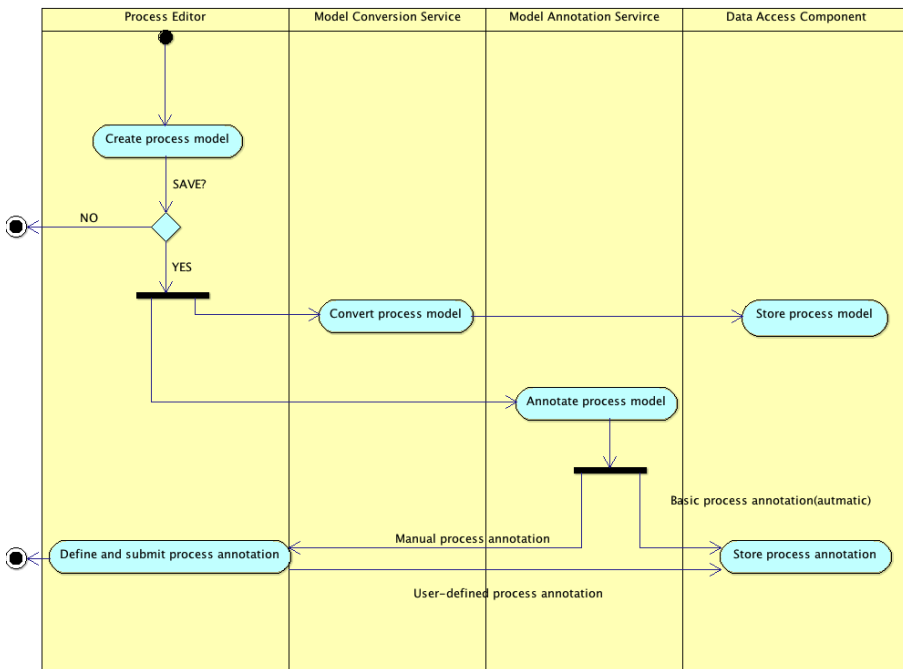


Figure 50: Populating the repository by creating a process model.

Figure 50 depicts the UML activity diagram for populating the repository by creating a process model in the repository. To populate the repository by creating a process model, a user uses the process editor accessed via the

Web to design a process model. A user may decide to store the process model or not. If the user decides to store the process model in the repository, the “Model Conversion Service” and “Model Annotation Service” are invoked. The Model Conversion Service, which is implemented by the process modelling engine, converts the model into XML format and stores it in the process model database. The Model Annotation Service, which is implemented by the Semantic Annotation Manager (SAM), automatically creates the basic process annotation and stores it in the process annotation database. Also, the annotation service provides an interface that lets the user perform manual process annotation based on the CPSAM annotation model, which is then stored in the process annotation database through the data access component.

9.3.4.2 Accessing Internal Repository

The search and retrieve process models are the most important tasks from the end-user perspective. Figure 51 depicts a UML activity diagram for searching a process model from the internal repository. The search is done based on the submitted query. By accessing the internal repository, the search service searches the process annotation database through the data access component.

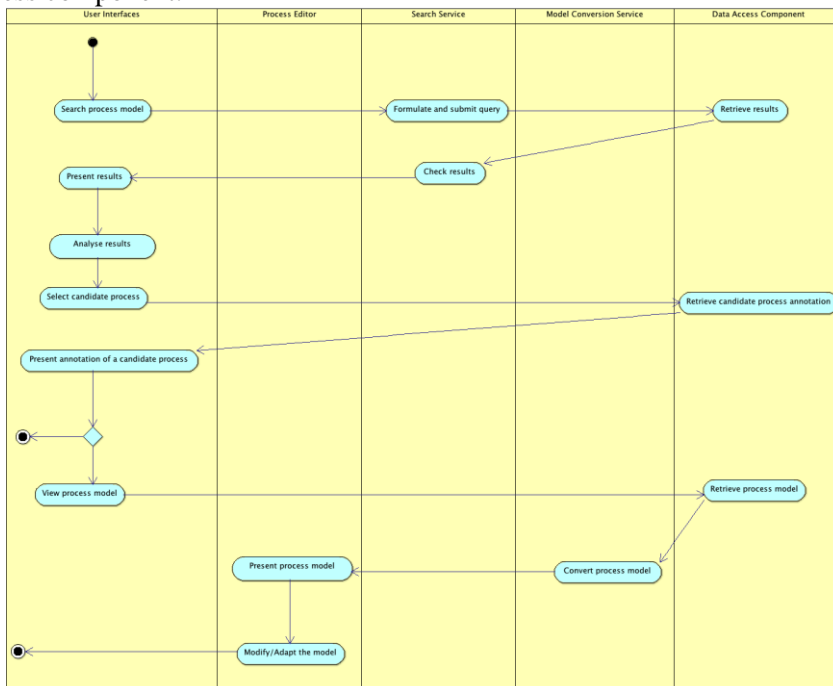


Figure 51: Search process model activity diagram.

It extracts a set of process models by using semantically based elements, and sends them to the user via the repository API. The user analyses the results

and selects the candidate process model from a list of displayed results. Once the candidate model is selected, its process annotation is retrieved and displayed to the user. The user may decide to quit or view the process model. To view the process model, the data access component retrieves the process model (in XML format) and triggers the Model Conversion Service, implemented by the process modelling engine, which converts the model into a graphical format in a respective modelling notation. The converted model is then displayed using a process editor on the Web browser. Finally, the user may modify and adapt the model for the specific use.

9.3.4.3 Accessing External Repositories

If there is no available or suitable process model in the internal repository, the search service sends a request to the process information mediator (PIM). To retrieve the process from external repositories the PMI performs a number of operations, including the following: (1) it accepts queries and rewrites them in a uniform query language, (2) it distributes and translates the query to queries with respective repository metadata elements, and (3) it returns a list of processes with the URIs, which points to the process models that meet the submitted query.

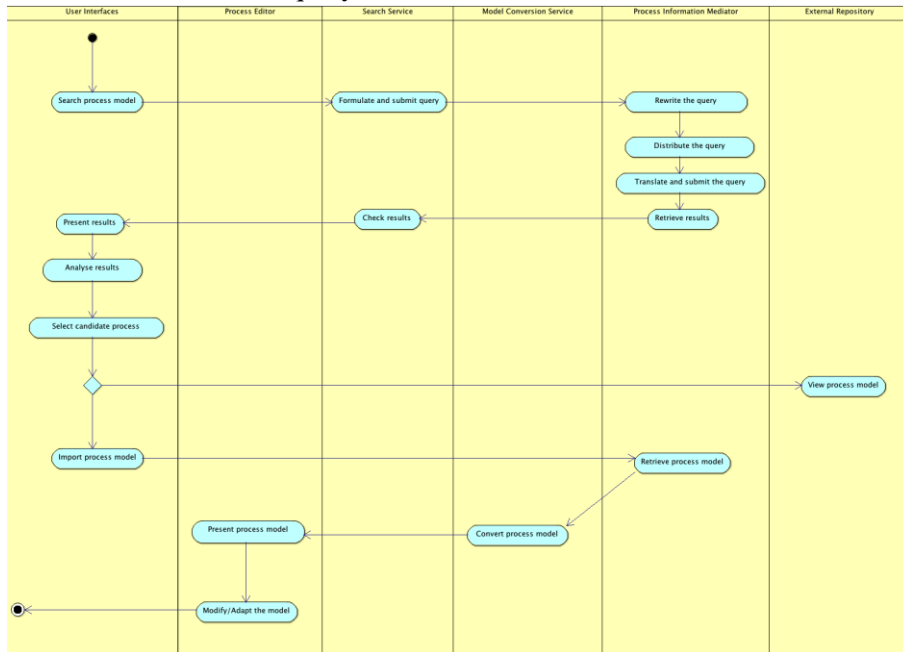


Figure 52: Accessing process models from external repositories.

During the search, when the results of the search are returned, the Model Annotation Service is triggered, which extracts and stores basic annotation of all the retrieved processes in the local repository; in this way the local repository can grow with a new set of process models. However, the actual

process models will not be stored locally unless the user decides to import them into the internal repository. Stored semantic annotation contains the URIs (Uniform Resource Identifiers) from which they can be accessed without querying through the mediator.

Once the search results are returned, the user analyses the results and selects the candidate process model from a list of displayed results. After selecting the candidate process, the user may decide to view the process model in the respective repository by refereeing the URIs, or import the process model into the repository from the external repository. The imported model is then converted by the Model Conversion Service and then presented to the user by the process model editor on the Web browser. At this point the user may modify and adapt the process model accordingly. The imported model is stored locally.

9.4 Architecture Evaluation

In this section we present the evaluation of the architecture as the evaluate artefact activity of a design science process. The evaluation is based on the analysis of the architecture against the requirements set out in Chapter 4.

One of the core requirements of the repository is to support standard modelling language (requirement R1). This architecture accomplishes that by provision of a *process modelling environment* that enables online creation of process models, storage and graphical representation of the stored models using standard notation. The process modelling environment includes three major components in the architecture: process model editor, process modelling engine and stencil set definitions. This also accomplishes the requirement for providing graphical representation of process models (requirement R3). The process modelling environment provides an online process editor for users to create new process models or edit existing ones. Therefore this also addresses the need to enable online modelling of process models (requirement 11).

The proposed architecture provides an advanced retrieval function based on semantic annotation. To accomplish that, the architecture provides a *semantic annotation manager* for semantically annotating process models before they are stored (requirements 7 and 8). Semantic annotations are based on the semantic annotation model (CPSAM), which includes a classification scheme and other meta information. The *repository engine* implements advanced process model retrieval through a search and navigation service. The use process annotation also enables a similarity search and comparison between process models (requirement 9).

The process annotations provided by the *semantic annotation manager* include definitions of process model relationships (requirement 10). During annotation the relationship between process models is created and stored as part of the process annotation.

Another core requirement is integration with external repositories to allow access to process models and growth of the repository (requirement 12). The architecture accomplishes this by introducing a *process information mediator* that enables interoperability and interconnection through wrappers, which transforms the query to respective repositories. The use of Web services enables the integration between different services offered by different system components of the repository.

To address the need for providing multiple process model granularities (requirement 5), the architecture includes a *process model abstraction engine*. The abstraction service component consists of the abstractor, which defines an abstraction algorithm, and the use case catalogue, which predefines use cases for which the model is intended. During model retrieval users can generate different process views that meet a desired need, through the abstraction service, which takes the desired use case definition from the use case catalogue sent with the initial process model to the abstractor. The abstractor generates a desired process model for the user.

Apart from providing a search and navigation service, the repository engine provides all fundamental functionalities that are common to any repository system such as version management (requirement 6) and configuration management (requirement 15). Therefore repository engine implements *version manager* and *configuration manager*.

To address the need for an access control mechanism (requirement 13), the repository architecture includes a *user manager component*, which provides authentication and authorization services. Users must be registered to exploit the search and retrieval features of the system along with a set of advanced functionalities, including a similarity check and model abstraction. Unregistered users can only perform simple search requests to find the desired process model.

The proposed repository is based on SOA, which enables extensibility (requirement 14) and scalability. The repository architecture provides extensibility by allowing existing repositories to integrate by just adding a new wrapper with mapping specification. In addition, through the use of Web services new functionalities can be added by simply advertising them to the repository. The semantic annotation service includes annotation edit, which allows new definition and annotation elements to be added. An object relational mapping mechanism enables additional definitions to be added in the process annotation schema.

Overall, the architecture fulfils the key requirements needed to implement the process model repository for process model reuse.

9.5 Related Work

There have been several attempts to design the architecture of process model repositories. One of the more recent and comprehensive works on the

process model repositories is that by La Rosa et al (2011a), who suggest an advanced process model repository (APROMORE).

APROMORE focuses on providing a place for advanced model-based analysis, filtering and consolidation. In their work they propose architecture for the process model repository, which consists of three layers: an enterprise layer, an intermediary layer and a basic layer. The enterprise layer is the front end of the repository, which hosts the repository manager. The basic layer encapsulates the business logic and data of traditional software architecture. The repository manager accesses both process models via the (de)canonization service – an intermediary adapter equipped with format conversion capabilities. Our architecture differs from their architecture in three main aspects: (1) apart from normal standard three-layered architecture, the architecture of the proposed repository introduces an interoperability layer that implements a process information mediator to enable exchange and sharing of process models between repositories; 2) to improve the efficiency of process model retrieval, the repository introduces a semantic annotation manager for semantically annotating process models in the repository. The search and navigation mechanisms of the repository are implemented based on the semantic annotation (3). In the proposed repository, process models can reside either locally or remotely. The central information stored in the repository is the process annotation, which includes the address of the process model. In addition, the use of semantic annotation enables process models to be stored and shared in any representation format such as modelling notations, textual format and images.

Yan et al (2012) is another work that proposes architecture of a process model repository. In their work they propose reference architecture for a business process model repository. Reference architectures are defined on a high level of abstraction as compared with concrete architectures, in this case the proposed repository architecture. In contrast, from concrete architectures, in general, reference architectures have a general nature and are designed to meet the attributes of all stakeholders. This makes it difficult to implement when you have a specific problem that is to be addressed.

9.6 Summary and Discussion

In this chapter we have proposed the architecture for an open and language-independent process model repository with an efficient retrieval system – a repository that is publicly open for change and growth by any potential user independently of the modelling language used and can comprise process models from existing process repositories. The main goal was not to suggest the best possible architecture, but to show that a good enough architecture can be designed based on the known architectural principles and knowledge sources. The principles and the knowledge sources were chosen based on

their fitness for the task at hand, the main requirement being that they can be integrated into a reasonable whole that can be used for developing a repository.

The proposed architecture is based on the requirements we introduced in Chapter 4. The architecture design is a work towards the development of an integrated and publicly open process model repository. This project is driven by the needs for an infrastructure to support reuse of process models.

The main contribution of this work is the architecture, which consists of four layers: the data, core, interoperability and presentation layer. The data layer stores all the process models, process annotation, process patterns, annotation model schema and user data. The core layer constitutes repository foundation services and repository service components, which implement the capabilities necessary for creating, manipulating and reusing the process knowledge. The interoperability layer encompasses the services that implement the mechanisms and an interface for external integration and accessing multiple heterogeneous data sources of the process model repositories. This integration is important to the success of reaching a critical mass of process models for reuse. It is one of the main features that are not offered in existing repositories. Another feature that is not offered in the existing repositories is that of structuring the process model repository provided by the semantic annotation manager. The main purpose of the semantic annotation manager is to semantically annotate process models in the repository for efficient retrieval of process models.

The proposed architecture has been evaluated against the requirement specification provided in Chapter 4. The results of the evaluation have indicated that the proposed architecture meets the requirements of the repository. Thus, we argue that the implementation of the proposed architecture will increase the probability of process model reuse.

10 Conclusion and Future Work

We conclude our work in this section with remarks on the accomplishments of the research goals with respect to the research contribution and limitations, and outline directions for future work.

10.1 Research Goals and Findings

In this research, we propose an open and language-independent process model repository with an efficient retrieval system to support the reuse of process models. The goals of this research included: (i) to elicit the requirements for a process model repository; (ii) to develop a semantic annotation model for annotating process models in the repository to facilitate searching of process models, navigating the repository and enhancing understanding of process models; (iii) to develop a business process relationship meta-model for identifying and defining the relationship between business processes in the repository; and (iv) to design the architecture of a process models repository for process model reuse.

Goal 1: To elicit the requirements for a process model repository

In order to address the shortcomings of existing process model repositories, in Chapter 4 we elicited a set of requirements for building such a repository from stakeholders (researchers and practitioners) and the literature, which formed the basis of the proposed process model repository. The requirements, provided in Chapter 4, were elicited by collecting initial results from an exploratory study (using semi-structured interviews) and validating the results through a confirmatory study (using a questionnaire). Based on the analysis of the exploratory study and the confirmatory study, a set of requirements for process model repositories was suggested by the stakeholders. Additional requirements were elicited from the literature through a systematic review approach. Finally, a specification and justification of each requirement was given. It should be noted that the presented requirements could be extended and adapted based on the primary purpose of the repository.

Goal 2: To develop and evaluate a semantic annotation model for semantically annotating process models to facilitate searching, navigation and understanding of process models

The main issue this research dealt with was the problem of locating and retrieving relevant process models. To address this problem, a *context-based process semantic annotation model (CPSAM)* was developed for annotating the process models – classifying and describing the process models in the repository. The annotation model development process constituted: the identification of potential annotation elements through a systematic literature review, validation of the elements through a confirmatory study, and construction of the model. The CPSAM, presented in Chapter 5, is based on well-established business frameworks, existing process classification schemes, organizational theories and other perspectives of a business process. The purpose of the annotation model is to facilitate searching for process models, navigating the repository and enhancing user understanding of process models.

To test and improve the performance of the annotation model, two empirical studies through controlled experiments were carried out. In the first study, the model was evaluated to test the annotation consistency and correctness. From the study, as discussed in Chapter 7, we learnt that the annotations of most of the CPSAM elements by different people are consistent (identical) and correct. In the second study, the model was evaluated to test the influence of the CPSAM on the searching, navigation and understanding of process models. From the study, as discussed in Chapter 7, we learnt that the CPSAM improves searching, navigation and understanding of process models.

Goal 3: To develop and evaluate a business process relationship meta-model for identifying and defining the relationship between business processes in the repository

The third issue addressed by this research is that related to identifying and defining relationships between business processes in the repository. The current definition of process model relationships is limited to directly related relationships. Such a relationship includes that of a process and its subprocesses, and the relationship of a process and a specialized process of the former. However, it is difficult to capture how all business processes are related in an enterprise. To address this problem, a business process relationship meta-model was developed for defining the relationship between business processes. The relationship meta-model development process constituted: the identification of potential relationship concepts, validation of the identified relationship concepts through a case study, and construction of the meta-model. The meta-model, presented in Chapter 8, is based on existing and well-established defined relationships and process-assets and asset-processes archetypes we have developed as a method to find

all processes that exist in an enterprise. The purpose of defining process relationships in the repository is to enable users to find related process models.

The meta-model was evaluated through an informed argument (Hevner et al, 2004) to test the extent to which it meets the established requirements. From the study, we learnt that the meta-model meets the established requirements.

Goal 4: To design the architecture of an open and language-independent process model repository with an efficient retrieval system to support reuse of process models

Based on the requirements provided in Chapter 4 and the results of the artefacts developed during the research project, the architecture of the process model repository was designed in Chapter 9. A Synthesis-based Software Architecture Design (SYNBAD) method was chosen for the design of the architecture. First, from the requirement specifications, we identified technical problems, which were then decomposed into subproblems. Secondly, a solution domain analysis was performed for each subproblem to identify architectural solutions. By adopting a Service-Oriented Architecture (SOA) as an architectural style the proposed architecture was created. The proposed architecture consists of five layers: the data layer, which stores all process knowledge; the business logic layer, which implements the core components of the repository; the service layer, which exposes the functionalities provided by the repository to the user through Web services and APIs; the presentation layer, which provides access of the repository services to the user; and lastly an interoperability layer, which implements a mechanism for repository integration with external or existing repositories.

The created architecture was then described based on the ISO/IEC/IEEE standards. For this, several viewpoints were used to describe the architecture. The data and information, functional and standard viewpoints have been used to describe the architecture. Finally, an informed argument was adopted to evaluate the proposed architecture against the specified requirements. From the analysis of the architecture, we can learn that the architecture fulfils the specified requirements.

In summary, the four research goals laid down have been achieved to a large extent; however, some few improvements may be required in the future. For the requirement elicitation, the study may need to consider other stakeholders, i.e. process owners and process users.

10.2 Research Contributions

Both the novelty of the artefact and the process of producing it constitute the contribution of a design science project (Hevner et al, 2004). The design

artefacts produced in this research include the sematic annotation model (CPSAM), process relationship meta-model and the architecture of the repository. Though not an artefact, we have also established a set of requirements for the process model repository as a contribution.

10.2.1 Contribution 1: Requirements for a business process model repository

A set of requirements for a process model repository is one of the contributions of this thesis. We have established a set of requirements that must be fulfilled by process model repositories in order to increase the probability of process model reuse. While some definitions of requirements for process model repositories existed, the elicitation of such requirements from a group of stakeholders is new. In addition, since process repositories can be designed for several different purposes, requirements for the repository to support reuse of process models are new. Therefore the set of requirements definitions provided in this thesis serves as an extension and validation of existing definitions of the requirements.

There have been a few attempts to establish requirements for business process model repositories. One of the early works is that of Shahzad et al (2009), who suggest a set of requirements for a business process model repository to support reuse of process models. Yan et al (2012) is another comprehensive work that defines a set of requirements for business process repositories. In both of these studies the requirements were elicited from reviewing existing process model repositories. To get a better understanding of the need and the problem, we have elicited the requirements from both the stakeholders and the literature including existing repositories. In addition the main focus of the requirements defined by Yan et al (2012) was to guide the design of the reference architecture of a business process repository. Reference architectures are more abstract, and therefore the defined requirements are also of a general nature for building a general-purpose repository. Our requirements are mainly focused on defining the requirements for designing a process model repository for supporting reuse of process models.

10.2.2 Contribution 2: A context-based process semantic annotation model

A context-based process annotation model is the core contribution of this thesis. The purpose of the model is to semantically annotate process models in the repository to facilitate retrieval of relevant process models (effective) with less effort (efficient) in a large collection. The novelty of the context-based process semantic annotation model (CPSAM) can be found in the conceptualization of the business framework (REA (Geerts & McCarthy,

2000; Dunn et al, 2005)), the conceptualization of existing process classification schemes (Porter's Value Chain (Porter, 2008), the Open-EDI framework (UN/CEFACT, 2003)) and the conceptualization of enterprise modelling concepts (Huat Lim et al, 1997; Fox et al, 1996). In Chapter 5, we described how it was developed, especially with insights and opinions from experts during the validation of annotation elements, which positively affirmed the concepts used. In addition, evaluation of the model in Chapter 7 confirmed that it is a versatile and applicable approach to improving process model retrieval. From a theoretical perspective, the annotation model provides an improved classification schemes for business process models. On the other hand, the practical impact of the semantic annotation model is to foster implementation of effective process retrieval systems for large collections of process models.

There have been several attempts to address semantic annotation of process models in the research literature. One of the more comprehensive approaches is that by Lin (2008), who suggests a semantic annotation framework to manage the semantic heterogeneity of process models. Born et al (2007) is another approach to address the issue of the semantic annotation of process models. In this work they proposes an approach for integrating semantics in modelling tools to support the graphical modelling of business processes with information derived from domain ontologies. Another work is that by the SUPER project (Wetzstein et al, 2007; SUPER, 2007), who proposes three main groups of ontologies for semantically annotating a business process: process, organization-related and domain-specific ontologies (Filipowska et al, 2009). Our work differs from these works in the respect that we focus on using annotated process models for the purpose of repository search and navigation. In contrast, other approaches have a wider scope and also intend to support the design of process models. As a consequence, these approaches require that process models be annotated with specific domain ontologies, typically tailored for the application and domain under consideration. CPSAM offers a lightweight approach, where well-known business and process frameworks are used as the basis for the annotation, which provides ease of annotation as well as ease of use. Nevertheless, the extensibility of CPSAM enables domain-specific notions and elements to be included if desired.

10.2.3 Contribution 3: A business process relationship meta-model

A business process relationship meta-model is another contribution of this thesis. The purpose of the meta-model is to enable users to identify and define the relationship between process models in the repository, which serves as the navigation mechanism. The novelty of the process relationship meta-model can be found in the conceptualization of the components of an enterprise (assets, sensor and processes) (Bider et al, 2011), and the

conceptualization of how business processes of an enterprise can be found (Bider et al, 2012). In Chapter 8, we described how it was developed, especially with insights and opinions from domain experts during the validation of the relationship concepts, which positively affirmed the concepts used. From a theoretical perspective, the relationship meta-model supports the understanding of relationships of business processes in the repository by reviewing how relationships are generally understood and supported in the repository. It also provides a holistic solution for representing the relationships in a uniform way. On the other hand, the practical impact of the relationship meta-model is to foster implementation of the navigation mechanism for effective process retrieval systems for large collections of process models.

There have been a few attempts to identifying and defining business process relationship in the repository. Malone et al (2003) are among those who can be noted for their pioneering of the concept of defining process model relationships in the repository. They define two types of relationship between business processes: (1) whole-part relationship, and (2) generalization-specialization relationship. Kurniawan et al (2012) extends the two types of relationship with (3) inter-operation relationship. The scope of these relationships is limited to directly related processes. They don't consider interdependencies between indirectly related processes in an organization. Given that the approach of using process model repositories to support reuse of process models is moving forward, significant effort in developing a holistic approach to identify and define the relationship between process models in the repository is necessary. The proposed relationship meta-model offers a holistic approach to identifying and defining process model relationships in the repository. The meta-model ties together the concepts from existing relationships and the process-assets and asset-processes archetypes with an understanding of how business processes are interconnected in the overall process architecture of an enterprise.

10.2.4 Contribution 4: Architecture for a business process model repository

The architecture of a process model repository is another contribution of this thesis. The purpose of designing the architecture of the process model repository is to bridge the existing gap by providing the basis for developing a repository system that will increase the probability of process model reuse. The main goal is not to suggest the best possible architecture, but to show that a good enough architecture can be designed based on the known architectural principles and knowledge sources. While we have followed some existing methods, the development, description and evaluation of the architecture serve as a methodological contribution to the body of knowledge. From a practical perspective, the architecture specified in this

thesis serves as a foundation for developing the actual repository to address the reuse problem.

There have been several attempts to design the architecture of process model repositories. One of the more recent and comprehensive architecture of a process model repository is that by La Rosa et al (2011a), which consists of three layers: an enterprise layer, an intermediary layer and a basic layer. Our architecture differs from their architecture in two main aspects: (1) apart from normal standard three-layered architecture, the proposed architecture introduces an interoperability layer that implements a process information mediator to enable exchange and sharing of process models between repositories; 2) In the proposed repository, process models can reside either locally or remotely and be accessed through the mediator. Yan et al (2012) is another work that proposes a reference architecture for a business process model repository. Reference architectures are defined on a high level of abstraction as compared with concrete architectures, in this case the proposed repository architecture. In contrast, from concrete architectures, in general, reference architectures have a general nature and are designed to meet the attributes of all stakeholders. This makes it difficult to implement when you have a specific problem that is to be addressed.

10.3 Limitations

This research set out to design an open and language-independent process model repository with an efficient retrieval system to support reuse of process models. Although we have succeeded in establishing the requirements, developing two conceptual models to improve the retrieval system and designing the architecture of such a repository, undeniable limitations exist that should be taken into consideration. These include:

- The evaluation of the semantic annotation model. While the model has been evaluated in two different studies, several limitations can be identified. (i) The first is the small size of the repository contents. Better results could be obtained if the evaluation was conducted with a large collection of process models. This was due to challenges of populating the repository with a large collection of process models. (ii) Secondly, the number of participants; a large number of participants are needed to generalize the findings. We had targeted a large number of participants to conduct the experiments, however some were not able to finish the required experimental tasks, as they required a considerable amount of time.
- The evaluation of the process relationship meta-model. While we have managed to perform a theoretical evaluation of the meta-model through an informed argument, the lack of a tool support limited the practical

and performance-based evaluation of the meta-model. The implementation of the meta-model is part of our future research.

- The evaluation of the repository architecture. While the proposed architecture is based on known architectural principles and knowledge sources, a practical evaluation of the architecture is necessary. This requires a complete implementation of all core components of the architecture, which is part of our future research.

10.4 Future Research Directions

In addition to continuing these aspects of the project there are several important areas of research that will be addressed by future project milestones. These include:

- Large-scale evaluation and refinement of the CPSAM. We plan to populate the repository with a large number of process models to enable further evaluation with a larger number of users, as well as a refinement of the semantic annotation model.
- Goal annotation template. In addition to that plan, our future work will also include the development of a formal template for goal formulation to enhance the goal annotation process.
- Implementation and practical evaluation of the process relationship meta-model. We plan to implement the meta-model in the repository to enable performance evaluation through experiments. This may involve the refinement of the meta-model depending on the feedback of the evaluation.
- The design and implementation of the process information mediator and integration with existing process model repositories. As part of the implementation of the architecture, the main focus will be to design and implement the mediator and integrate with some of the existing repositories for sharing process models.

Appendices

A.1. Survey of Existing Process Model Repositories

The survey reported in this appendix is a published work in Elias and Johannesson (2012b).

The survey of existing repositories is based on publications in academic as well as trade journals and conferences. In order to identify process model repositories, we searched for relevant journal and conference publications by searching Google Scholar using the four keyword phrases “business process model repository”, “process model repository”, “business process repository” and “process repository”. The results from these searches were narrowed down to publications fulfilling the following criteria: (i) the title of the publication explicitly includes or implicitly refers to the area of interest, i.e. business process model repositories, (ii) the publication describes or proposes a business process model repository, and (iii) the publication has a citation score of at least 5 in Google Scholar.

Based on the above criteria, 26 publications were selected, and ten repositories of business process models were identified: (i) MIT Process Handbook (MIT process handbook project, 2001; Malone et al, 2003), (ii) Phios process repository for supply chain (SCOR) (Supply Chain Council, 2003), (iii) IBM Process Repository (IBM Corporation, 2004), (iv) IBM-BPEL Repository (Jussi et al, 2006), (v) Semantic Business Process Repository (Ma et al, 2007), (vi) Oryx (Decker et al, 2008; Oryx, 2008), (vii) SAP Process Repository (SAP AG, 2007), (viii) Prosero Repository (Elhadad et al, 2008), (ix) RepoX Repository (John, 2001), (x) Advanced Process Model Repository (La Rosa et al, 2011a).

In this section, we briefly review the identified repositories. Aspects for reviewing these repositories have been derived from the requirements suggested in the previous section. The aspects include *openness*, *standard modelling language support* (the repository should be able to store process models in at least one standard process modelling language), *representation* (process models in the repository should be represented in both graphical and textual form), *domain independence* (the repository should allow process models to be stored regardless of their domain – storing both domain-specific and generic process models), *business model inclusion* (the repository should store both business and process models), *classification scheme* (process models in the repository should be categorized based on

widely accepted classification schemes to facilitate navigation), *goal inclusion* (a process model should be annotated with information that can facilitate searching, navigating and interpreting process models) and *versioning* (the repository should allow multiple versions of a process model to be maintained). Openness and goal inclusion are the additional aspects. Openness is the public availability of the repository to its potential users without any proprietary constraints – the ability of users to add, update, delete or retrieve process models without any prior legal permission. The cost element has not been considered. Another additional aspect is goal inclusion. The relationship between *goal inclusion* and requirement 7 can be traced back from the exploratory study. An aspect related to process granularity levels as a requirement is not included. This is because granularity is an intentional process design, however at the repository level, this is met if the repository supports multiple versions of the same process.

MIT Process Handbook

Openness. The MIT process handbook is a proprietary repository that provides a knowledge base of process descriptions (MIT process handbook project, 2001; Malone et al, 2003). Therefore, the repository can only be extended and enhanced by its owners.

Standard modelling language. In the MIT repository, business processes are described in natural language, and no standard modelling notation is supported.

Representation. With respect to process representation form, the MIT process handbook only describes business processes in textual form.

Domain independence. The repository is not restricted to any specific domain.

Business model inclusion. The repository does not include business models.

Classification scheme. The process classification scheme adopted by the MIT process handbook is based on two dimensions: generalization-specialization and composition-decomposition, where each process in the repository can be viewed. While browsing of processes is based on the two dimensions of the classification scheme, the repository only provides a keyword-based search for finding processes.

Goal inclusion. Process models stored in the repository are not related to goals.

Versioning. The repository does not manage multiple versions of process models.

Phios Process Repository for Supply Chain (SCOR)

Openness. SCOR (Supply Chain Council, 2003) is another proprietary repository similar to the MIT process handbook. It provides a knowledge base of process descriptions related to supply chain management.

Standard modelling language. Like the MIT process handbook, business processes are described in natural language, and no other modelling notation is supported.

Representation. Like the MIT repository, processes are represented in textual form.

Domain independence. SCOR is restricted to supply chain management processes.

Business model inclusion. The repository does not include business models.

Classification scheme. Apart from the two dimension classification scheme, adopted by the MIT process handbook, processes in SCOR are further classified based on four verbs: create, destroy, modify and preserve. In addition, processes are organized around five management root processes: plan, source, make, deliver and return.

Goal inclusion. Process models stored in the repository are not related to goals.

Versioning. The repository does not manage multiple versions of process models.

IBM Process Repository (IBM PR)

Openness. The IBM process repository is proprietary to IBM (IBM Corporation, 2004).

Standard modelling language. The notation used is not standard but specific to the repository.

Representation. Process models are represented in both textual and graphical form with the aim of providing an explicit control flow.

Domain independence. IBM PR is restricted only to e-commerce business processes.

Business model inclusion. The repository does not include business models.

Classification scheme. Processes are classified into five major groups: B2B direct, consumer direct, demand chain, hosting and supply chain. In each group, processes are further classified into three subgroups: direct admin processes, direct starter stores and direct solution.

Goal inclusion. The repository includes the objectives of each process, but not business goals. In contrast to an objective, a goal tends to be longer term, general (rather than specific), qualitative (rather than quantitative) and ongoing.

Versioning. The repository does not handle multiple versions of process models.

IBM-BPEL Repository

Openness. The IBM-BPEL repository is a proprietary repository from IBM for storing and retrieving process models expressed in the Business Process Execution Language (BPEL) format (Jussi, 2004; Jussi et al, 2006).

Standard modelling language. The repository only supports BPEL.

Representation. Processes are represented using a BPEL XML format and stored internally as objects in an ECLIPSE repository.

Domain independence. The repository is not restricted to any specific domain.

Business model inclusion. The repository does not include business models.

Classification scheme. The repository does not include a process classification scheme.

Goal inclusion. Process models stored in the repository are not related to goals.

Versioning. The repository does not handle multiple versions of process models.

Semantic Business Process Repository (SBPR)

Openness. SBPR (Ma et al, 2007) is a non-proprietary repository for storing and managing semantic business process models in SBPM.

Standard modelling language. The repository supports Business Process Modelling Ontology (BPMO), sBPEL, sBPMN and sEPC, which are the ontological versions of BPEL, Business Process Modelling Notation (BPMN) and Event-Driven Process Chains (EPCs) [25].

Representation. In SBPR, processes are described in graphical form.

Domain independence. The repository is not restricted to any specific domain.

Business model inclusion. The repository does not include business models.

Classification scheme. The repository does not include a process classification scheme.

Goal inclusion. Process models stored in the repository are not related to goals.

Versioning. The repository supports (manages) multiple versions of process models.

Oryx

Openness. Oryx (Decker et al, 2008; Oryx, 2008) is a non-proprietary repository that provides a Web-based process modelling tool to enable users to create, store and update process models online.

Standard modelling language. Oryx supports several process modelling notations, including BPMN, Petri nets and EPC.

Representation. Business processes are represented in graphical form. Process models are stored in a database and externally represented in RDF format.

Domain independence. The repository is not restricted to any specific domain.

Business model inclusion. The repository does not include business models.

Classification scheme. The repository does not include a process classification scheme.

Goal inclusion. Process models stored in the repository are not related to goals.

Versioning. The repository does not manage multiple versions of process models.

SAP Business Map

Openness. The SAP business map (SAP AG, 2007) is a proprietary repository.

Standard modelling language. The notation used is not standard but specific to the repository.

Representation. The repository provides process models in a graphical form, which includes its purpose, prerequisites and activity flows.

Domain independence. The repository is limited to its application products.

Business model inclusion. The repository does not include business models.

Classification scheme. Processes in the SAP business map are classified into eight major business scenarios: interaction centre for automation, make-to-order production in supply chain management (SCM), order-to-delivery, release processing, supplier-managed inventory, Radio Frequency Identification (RFID)-enabled returnable transport items, Web-based supplier Kanban and dealer business management.

Goal inclusion. The purpose of each process model is included but not business goals.

Versioning. The repository does not handle multiple versions of process models.

Prosero

Openness. Prosero is an SOA-based semantic repository of business processes and Web services (Elhadad et al, 2008) meant to be used by an enterprise and its customers.

Standard modelling language. The repository only supports the BPMN notation. For execution, the repository provides a BPEL generator that transforms process models from BPMN into BPEL.

Representation. Business processes are represented in graphical form.

Domain independence. The repository is not restricted to any specific domain.

Business model inclusion. The repository does not include business models.

Classification scheme. The repository does not include a process classification scheme.

Goal inclusion. Process models stored in the repository are not related to goals.

Versioning. The repository does not handle multiple versions of process models.

RepoX

Openness. RepoX (John, 2001), an XML repository management tool, is a client-server model (not publicly open) repository developed in the METEOR Workflow System environment for the purpose of managing XML-based metadata.

Standard modelling language. Definitions of workflow processes are stored as metadata in the form of XML documents.

Representation. Business processes are represented in graphical form.

Domain independence. The repository is not restricted to any specific domain.

Business model inclusion. The repository does not include business models.

Classification scheme. The repository does not include a process classification scheme.

Goal inclusion. Process models are not related to business goals.

Versioning. RepoX supports and manages multiple versions of process models.

Advanced Process Model Repository (APROMORE)

Openness. APROMORE (La Rosa et al, 2011a) is an SOA-based (non-proprietary) repository that offers a rich set of features to maintain, analyse and exploit the content of process models.

Standard modelling language. In APROMORE, business processes are described in a common format (called “canonical format”).

Representation. In APROMORE, processes are represented in graphical form.

Domain independence. The repository is not restricted to any specific domain.

Business model inclusion. The repository does not include business models.

Classification scheme. The repository does not include a process classification scheme.

Goal inclusion. Process models stored in the repository are not related to goals.

Versioning. The repository supports (manages) multiple versions of process models.

Analysis of Existing Process Model Repositories

Based on the above review and the summary shown in Table 20, we analyse the surveyed repositories by identifying challenges to be addressed. The main challenges are the following:

- *Openness*: Most of the repositories, except APROMORE, SBPR and Oryx, are proprietary, i.e. they are the intellectual property of some organizations. The repositories do not allow users outside these organizations to add, update, delete or retrieve process models, without prior legal permission. This lack of openness can impede the acceptance and consequent use of the repositories, thereby making it more difficult to achieve a critical mass of process models available for reuse.
- *Standard modelling language*: While it is not necessary to support multiple languages in order to provide reusable process models (Shahzad et al, 2010), the support of at least one standard modelling notation is necessary. Process models in some of the repositories, such as MIT, SCOR, IBM PR and SAP, are given in non-standard modelling notations, which makes it difficult to transform them into executable models or to users' modelling notations of interest for reuse.
- *Domain independence*: Some of the repositories, such as IBM PR, SCOR and SAP, have a restricted scope, as they are limited to certain domains. IBM PR is restricted to e-commerce, SCOR is restricted to supply chain management and SAP is restricted to application products. Restricting the repository to certain domains affects the growth of the repository and the reusability of models between business domains becomes restricted.

Table 20: Repositories review summary

Requirements \ Repositories	MIT	SCOR	IBM-PR	IBM-BPEL	SBPR	Oryx	SAP	Prosero	Repox	APROMORE
Openness	-	-	-	-	+	+	-	-	-	+
Standard modelling language support	-	-	-	+	+	+	-	-	+	+
Domain independence	-	-	-	+	+	+	-	+	+	+
Process representation	T	T	G&T	G	G	G	G&T	G	G	G
Business model inclusion	-	-	-	-	-	-	-	-	-	-
Versioning	-	-	-	-	+	-	-	-	+	+
Goal inclusion	-	-	-	-	-	-	-	-	-	-
Classification scheme	+	+	+	-	-	-	+	-	-	-

In the table above T stands for textual representation and G for graphical representation.

- *Process representation*: In most of the repositories, processes are represented in either graphical or textual form and not both. Only IBM

PR and SAP provide both graphical and textual representation of process models, however they both use non-standard notation and their textual format is not well structured to capture important aspects of a business process. This affects understanding of process models by users, which affects reusability.

- *Business model inclusion*: None of the surveyed repositories include business models; therefore a high-level view of the business activities performed by an enterprise is not given. This makes it difficult for users to get a better understanding of process models that could meet their business requirements.
- *Versioning*: Most of the repositories offer a single process model for certain business process scenarios. The repositories, except for SBPR, RepoX and APROMORE, do not provide support to manage multiple versions of process models for the same business process. This lack of multiple-version support may lead to a loss of process knowledge if new ones replace existing models.
- *Goal inclusion*: In most of the repositories, the process models are not related to goals. This makes it difficult for users to gain an understanding of the business goals that are realized by a certain process. As achieving business goals is the purpose of a process, the lack of explicit goal representation also makes it more difficult to understand the process models themselves.
- *Classification scheme*: Searching and navigating across repositories is often a complex and time-consuming task, making it difficult for users to find relevant process models. One reason for this is that most repositories offer their own proprietary process classification schemes instead of utilizing more standard and well-established schemes. As a consequence, users need to understand and learn these proprietary schemes, which makes searching and navigating more demanding.

Some of these challenges are related to the intended use of a repository. In particular, a restriction to a specific domain is typically an intentional design decision. Other challenges are due to economic and organizational factors, such as the decision of whether to make a repository proprietary or open. However, the challenge of facilitating search and navigation is common to all repositories. This is because of the lack of standard and well-established classification schemes. Another major challenge is the lack of an efficient version management technique for business processes stored in the repositories. In addition to classification schemes and versioning, another challenge is the difficulty of identifying and understanding business processes that meet users' business needs. This is because stored process models are not well described to help users identify process models that might meet their needs.

A.2. Requirements Elicitation Questionnaire

Goal: Collect requirements for a process repository by using an exploratory study.

Personal data

1. Name: _____
 2. Position: _____
 3. Company: _____
- Etc.

1. Do you build process models and if so, do you build them from scratch?
 - a. If yes,
 - i. If you face any modelling problems, please explain them briefly.
 - ii. What are the specific reasons that you don't reuse process models?
 - iii. Do you want to reuse process models?
 - iv. Do you plan to reuse process models in the future?
 - b. If no,
 - i. How do you reuse process models? // if you don't build them from scratch, then you probably reuse.
 - ii. Are you using any repository for process models right now?
 - iii. How do you store and retrieve the process models to be reused?
 - iv. What are the benefits of reusing process models?
2. Have you used different languages for process modelling?
 - a. If yes,
 - i. Have you found process models (written in different languages) easy to compare and relate?
 1. If they are difficult to relate and compare, does it hinder reuse?
 2. Do you think a repository can bridge the gap between process models written in different languages?
 - a. If yes, how?
 - b. If no,
 - i. Which language do you use for process modelling?
3. If you have used any process model repositories, please state their names.
4. What are, in your opinion, the most important properties of a process repository that support reuse of process models? //explain
5. Goal alignment means that the goals of a process are stated within the process description. Will goal alignment affect process model reuse?
 - a. If yes, how?
 - b. Are your process models aligned with business goals?

6. Customizability is the ability to make changes to process models stored in a repository. Do you think customizability affects reuse of process models?
 - a. If yes,
 - i. In a repository, how can process models be customized?
 - ii. How can multiple users store and retrieve a process model that they have customized?
7. In your company, are all the process models developed at the same level of detail?
 - a. If the level of detail is different, do you find any problems in relating, comparing and reusing them?
 - b. How should process models that differ in their level of detail be stored in a process repository?
8. There are two main user groups: IT users and business users. Communication gaps between business and IT users are often problems.
 - a. Do you experience communication gaps between business and IT users in your organization?
 - i. If yes,
 1. Can a repository play a role in reducing the gap between the two user groups?
 2. Should these users share a common repository of process models or should they work on separate repositories?
 - b. If no, give an example that you know of a process model that is understandable both by IT and business users.
9. A classification scheme provides bases for grouping process models. Please suggest some classification schemes for a process repository that can be used for enhancing searching, navigating and reusing process models.
10. What role can a process model repository play? // we discussed one role of a repository.
11. Any other requirements/suggestions for a process model repository?

Any other feedback.

A.3. Requirements Confirmatory Study Instruments

Purpose: To validate the requirements for a process model repository elicited from an exploratory study.

Position _____ Affiliation _____

Profession: Researcher/Practitioner/Both _____ Area of Expertise _____

	5	4	3	2	1	Reasoning/ Comments (if any)
Questions on Process models						
1. A domain expert can understand a process related to its domain written in any process modelling language.	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree	
2. Experts of a process modelling language with common knowledge* of a domain can understand any process model in that domain written in a language of his/her expertise.						
3. Reuse of process models can simplify the work of modelling business processes, improve modelling efficiency and reduce the cost of modelling business processes.						

4. A process model written in one process modelling language is difficult to reuse for users of another language.					
5. Domain-independent process models** can be reused for modelling specific business processes in an enterprise.					
6. A graphical representation of a business process is easier to understand than a textual representation.					
7. Communication gaps between business experts and IT designers often cause problems in understanding process models.					
8. A process model repository can help reduce the gap between business experts and IT designers.					
Questions on process model repository					
9. A process model repository can play an important role in reusing process models.					
10. A repository can support reuse even if only fundamental elements (activities, agents, control flow) of process models are stored, i.e. composite tasks, intermediate events etc. are omitted.					
11. In a repository, it is useful to represent the same process using several process models with different levels of detail.					
12. A repository should maintain multiple versions of all its process models.					

* *Common knowledge*: the facts generally known to everyone.

** *Domain processes*: a process model that is not specific to a particular industry.

Comments on the questionnaire (if any).

Thank you for your participation!

Khurram Shahzad, Mturi Elias, Paul Johannesson
Information Systems Laboratory (Syslab),
Department of Computer and System Sciences (DSV),
Royal Institute of Technology (KTH)/Stockholm University (SU), Sweden.

A.4. Validation of Potential Annotation Elements

Purpose: To validate the potential annotation elements from the literature review.

Position _____ Affiliation _____

Profession: Researcher/Practitioner/Both _____ Area of Expertise _____

	5	4	3	2	1	Reasoning/Comments (if any)
	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree	
The following are important to search, navigate and interpret process models in a repository.						
<i>Process description (textual description)</i>						
<i>Business context</i>						
<i>Business goals</i>						
<i>Domain-specific classification schemes (like SCOR classification for Supply Chain etc.)</i>						
<i>Generic classification schemes (like classification based on Open EDI and Porter's Value Chain etc.)</i>						
<i>Properties of processes (fast, cost reduction etc.)</i>						

Resources (services, goods, money etc.)

Actors (customer, supplier etc.)

Relationship between processes

Comments on the questionnaire (if any).

Thank you for your participation!

Mturi Elias, Paul Johannesson
Information Systems Laboratory (Syslab),
Department of Computer and System Sciences (DSV),
Royal Institute of Technology (KTH)/Stockholm University (SU), Sweden.

A.5. CPSAM Evaluation Instruments

Experiment Tasks

Task 1: Process Understandability (Correct Answers & Time)

Question: For each of the process models provide the following:

1. Briefly explain the purpose of the process model
2. Identify the resource being consumed or received during process execution
3. Identify the main actor and the role played by the actor
4. Identify activities that may not be executed at all

Task 2: Searching Process Models

Question: For each of the search requests perform both keyword-based search and annotation-based search.

1. Find an order-to-cash process model for the delivery of services (e.g. consultancy services) to the customer
2. Find the process model for disbursing credit to the customer (loan applicant) by the creditor (bank)
3. Find all process models related to procurement
4. Find all process models related to recruitment of employees
5. Locate a process model that allows strategic personnel (top management) to carry out marketing analysis

Task 3: Navigating the Repository

Question: For each of the search requests perform both alphabetical-based navigation and annotation-based navigation.

1. Find a process model for carrying out service and maintenance of an asset after receiving the fault report
2. Find a procurement process model for selecting a supplier by requesting and evaluating quotations from several suppliers
3. Find a process model for disbursing credit to the customer (approved loan applicant) by the creditor (bank)
4. Find a human resource process model for getting a new employee ready for work
5. Find a process model that allows the creditor to assess the credit risk and provide assessment information

Task 4: Filling in Post-Task Survey Questionnaire

POST-TASK SURVEY QUESTIONNAIRE	5	4	3	2	1	Comments
	Strongly Agree	Agree	Not Sure	Disagree	Strongly Disagree	
1. I think the annotations have improved the searching of process models in the repository						
2. I found navigating the process model repository based on CPSAM elements to be more efficient than the alphabetical approach						
3. I found the annotations to be helpful for understanding and interpreting process models						
4. It was easy for me to search the process models by using annotation-based search						
5. It was easy for me to navigate the repository by annotation-based navigation						
6. If I have to search a process in the repository in the future I will use annotation-based search						
7. If I have to navigate the repository in the future I will use an annotation-based approach						
8. If I am involved in building the repository for process models I would recommend the CPSAM model						

A.6. Case Study Interview Guide

Case Study Interview Guide

Purpose: To investigate the business processes (and their relationship) needed to support the education service delivery in a higher education institution.

Position: _____ Affiliation _____

Profession: _____ Area of Expertise _____

Introduction

The primary business of most higher education institutions (HEIs) is education service delivery, research and consultancy. The focus of this study is on the education service delivery, which includes teaching, examination and graduation processes. Our assumption is that these processes are well known, visible and documented. Therefore we intend to investigate other processes that are vital to the successful implementation of the education service delivery as the main business process.

Education

1. Teaching and learning involve all processes linked to delivering knowledge to students. They include curriculum development, instructional design, teaching, examining and graduation. Are these processes designed and documented?
2. The above processes require the HEI to utilize some resources to ensure the effective delivery of knowledge. Do you agree that human capital, educational materials, finances, infrastructure, libraries and partners are the major resources that need to be in place for the education service delivery?
3. What are the other resources that need to be available for the education service delivery?
4. How are the educational materials (i.e. books, instructional materials) managed?

What business processes does the institute utilize for:

- a. Acquiring them,

- b. Maintaining and making them available at the right time and point (i.e. the students)?
- 5. Do you agree that you need some plans, policies and strategies for proper management of teaching and learning? If yes, what are these plans, policies and strategies?
- 6. Who are your primary customers? How does your institute attract and get new customers (i.e. students, companies)? What do you do to keep your current customers?
- 7. Who are your business partners? How do you find new partners and what process does the institute utilize to ensure your relationship is well maintained?

Human Capital

- 8. How is the human capital (lecturers, administrators, support staff, etc.) for your institute managed? What business processes does the institute utilize:
 - a. For getting (acquiring) lecturers, administrators and support staff)?
 - i. What resources are needed for each of the mentioned processes?
 - b. To ensure that the employed human resource is available in a good working condition?
 - i. What resources are needed for each of the mentioned processes?
 - c. For separation with the employed human resource,
 - i. What resources are needed for each of the mentioned processes?
- 9. Do you agree that you need some plans, policies and strategies for proper management of the human capital? If yes, what are these plans, policies and strategies?

Financial Resources

- 10. What are the main sources of financial resources for the institute?
- 11. How is the finance for the institute managed?
What business processes does the institute utilize?
 - a. For getting (acquiring) finances,

- i. What resources are needed for each of the mentioned processes?
 - b. To ensure that the finances are properly managed,
 - i. What resources are needed for each of the mentioned processes?
- 12. Do you agree that you need some plans, policies and strategies for proper management of the financial resources? If yes, what are these plans, policies and strategies?

Technical and Information Infrastructure

- 13. What are the technical and information infrastructures the institute possesses?
- 14. How are these infrastructures managed?
What business processes does the institute utilize for:
 - a. Acquiring the infrastructure?
 - b. Maintaining and making the infrastructure available at the right time?
- 15. Do you agree that you need some plans, policies and strategies for proper management of the infrastructure? If yes, what are these plans, policies and strategies?

Comment (if any).

Thank you for your participation!

Mturi Elias, Ilia Bider, Paul Johannesson

Information Systems Laboratory (Syslab), Department of Computer and System Sciences (DSV), Stockholm University (SU), Sweden.

References

- Agarwal, R., De, P. & Sinha, A. P. (1999) Comprehending object and process models: An empirical study. *IEEE Transactions on Software Engineering*, 25 (4), 541–556.
- Aldin, L. (2010) Semantic Discovery And Reuse Of Business Process Patterns. *School of Information Systems, Computing and Mathematics*. London, Brunel University, pp. 245.
- Aldin, L. & de Cesare, S. (2011) A literature review on business process modelling: new frontiers of reusability. *Enterprise Information Systems*, 5 (3), 359–383.
- Alexander, C., Ishikawa, S. & Silverstein, M. (1977) *A Pattern Language: Towns, Buildings, Construction (Center for Environmental Structure Series)*, New York, Oxford University Press.
- Ali, F. M. & Du, W. (2004) Toward reuse of object-oriented software design models. *Information and Software Technology*, 46 (8), 499–517.
- American Productivity and Quality Center. (2001) *APQC – Process Classification Framework*. [Online] Available from: <http://www.apqc.org/process-classification-framework>. [Accessed 2nd September, 2014.]
- Andersson, B., Bider, I., Johannesson, P. & Perjons, E. (2005) Towards a formal definition of goal-oriented business process patterns. *Business Process Management Journal*, 11 (6), 650–662.
- Andrews, T., Curbera, F., Dholakia, H., Goland, Y., Klein, J., Leymann, F., Liu, K., Roller, D., Smith, D., & Thatte, S. (2003) Business process execution language for web services, Version 1.1 Specification, BEA Systems, IBM Corporation, Microsoft Corporation, SAP AG, Siebel Systems.
- Anthony, R. N. (1995) *Planning and Control Systems: A Framework for Analysis*. Cambridge, MA, Harvard Business School.
- Anthony, R. N., Dearden, J. & Bedford, N. M. (1995) *Management Control Systems*. Homewood, IL, Irwin.
- Arkin, A. (2003) *Business process modeling language*. [Online] Available from: <http://www.omg.org/bpmn/Documents/BPML-2003.pdf>. [Accessed 28th April, 2014]
- Armistead, C., Harrison, A. & Rowlands, P. (1995) Business process re-engineering: lessons from operations management. *International Journal of Operations & Production Management*, 15 (12), 46–58.
- Bae, H., Cho, E. & Bae, J. (2007) A version management of business process models in BPMS. In: Chang, K., Wang, W., Chen, L., Ellis, C., Hsu, C., Tsoi, A., & Wang, H. (eds.) *Advances in Web and Network Technologies, and Information Management*. Berlin / Heidelberg, Springer, pp. 534–539.
- Barjis, J. (2011) CPI modeling: Collaborative, participative, interactive modeling. In: Jain, S., Creasey, R.R., Himmelspach, J., White, K.P., Fu, M. (eds.),

- Proceedings of the 2011 Winter Simulation Conference*. Piscataway, NJ, pp. 3099–3108.
- Barros, A. & Borger, E. (2005) A compositional framework for service interaction patterns and interaction flows. In: Lau, K.-K. & Banach, R. (eds.) *Formal Methods and Software Engineering*. Berlin Heidelberg, Springer, pp. 5–35.
- Barros, A., Dumas, M. & ter Hofstede, A. (2005a) *Service Intercation Patterns*. [Online] Available from: <http://math.ut.ee/~dumas/ServiceInteractionPatterns/>. [Accessed 2nd September, 2014.]
- Barros, A., Dumas, M. & ter Hofstede, A. M. (2005b) Service interaction patterns. In: van der Aalst, W. P., Benatallah, B., Casati, F. & Curbera, F. (eds.) *Business Process Management*. Berlin Heidelberg, Springer, pp. 302–318.
- Barros, O. (2007) Business process patterns and frameworks: Reusing knowledge in process innovation. *Business Process Management Journal*, 13 (1), 47–69.
- Bauer, M. I. & Johnson-Laird, P. N. (1993) How diagrams can improve reasoning. *Psychological Science*, 4 (6), 372–378.
- Becker, M. & Klingner, S. (2014) A criteria catalogue for evaluating business process pattern approaches. In: Bider, I., Gaaloul, K., Krogstie, J., Nurcan, S., Proper, H., Schmidt, R., & Soffer, P. (eds.) *Enterprise, Business-Process and Information Systems Modeling*. Berlin Heidelberg, Springer, pp. 257–271.
- Bergholtz, M., Jayaweera, P., Johannesson, P. & Wohed, P. (2003) Process models and business models – a unified framework. In: *Proceedings of the ER Workshops*. Berlin / Heidelberg, Springer, pp. 364–377.
- Bernstein, P. (1997) Repositories and object oriented databases. In: Dittrich, K. & Geppert, A. (eds.) *Datenbanksysteme in Buro, Technik und Wissenschaft (Proceedings of BTW Conference)*. Berlin Heidelberg, Springer, pp. 34–46.
- Bernstein, P. A. & Dayal, U. (1994) An overview of repository technology. In: *Proceedings of the 20th International Conference on Very Large Data Bases*. San Francisco, Morgan Kaufmann Publishers Inc., pp. 705–713.
- Bider, I., Perjons, E. & Elias, M. (2012) Untangling the dynamic structure of an enterprise by applying a fractal approach to business processes. In: Sandkuhl, K., Seigerroth, U. & Stirna, J. (eds.) *The Practice of Enterprise Modeling*. Berlin Heidelberg, Springer, pp. 61–75.
- Bider, I., Bellinger, G. & Perjons, E. (2011) Modeling an agile enterprise: Reconciling systems and process thinking. In: Johannesson, P., Krogstie, J. & Opdahl, A. (eds.) *The Practice of Enterprise Modeling*. Berlin Heidelberg, Springer, pp. 238–252.
- Bider, I. & Perjons, E. (2009) Evaluating adequacy of business process modeling approaches. *Handbook of Research on Complex Dynamic Process Management: Techniques for Adaptability in Turbulent Environments*. Hershey, New York, Business Science Reference, pp. 79–102.
- Born, M., Dörr, F. & Weber, I. (2007) User-friendly semantic annotation in business process modeling. In: Weske, M., Hacid, M.-S. & Godart, C. (eds.) *Web Information Systems Engineering (WISE 2007) Workshops*. Berlin / Heidelberg, Springer, pp. 260–271.
- Brockmans, S., Ehrig, M., Koschmider, A., Oberweis, A. & Studer, R. (2006) Semantic alignment of business processes. In: Manolopoulos, Y., Filipe, J., Constantopoulos, P., & Cordeiro, J. (eds.) *Proceedings of the Eighth*

- International Conference on Enterprise Information Systems (ICEIS 2006)*. Paphos, Cyprus : INSTICC Press, pp. 191–196.
- Bullinger, A. (2008) *Innovation and Ontologies: structuring the early stages of innovation management*. Melbourne, Gabler Verlag.
- Cardoso, J., Mendling, J., Neumann, G. & Reijers, H. A. (2006) A discourse on complexity of process models. In: Eder, J. & Dustdar, S. (eds.) *Business Process Management Workshops*. Berlin Heidelberg, Springer, pp. 117–128.
- Chang, J. F. (2005) *Business Process Management Systems: Strategy and Implementation*. New York, Auerbach Publications.
- Chavan, C. R. & Lal, B. (2012). Analysis of Business Processes and Modeling Approach to Business Process Re-Engineering. *International Journal of Computer Science and Information Technologies*, 3 (4), 4731 – 4737
- Constantopoulos, P., Jarke, M., Mylopoulos, J. & Vassiliou, Y. (1995) The software information base: A server for reuse. *The VLDB Journal*, 4 (1), 1–43.
- Curtis, B., Kellner, M. I. & Over, J. (1992) Process modeling. *Communications of the ACM*, 35 (9), 75–90.
- Danesh, A. & Kock, N. (2005) An experimental study of process representation approaches and their impact on perceived modeling quality and redesign success. *Business Process Management Journal*, 11 (6), 724–735.
- Davenport, T. H. (1993) *Process Innovation: Reengineering Work Through Information Technology*. Boston, Harvard Business School Press.
- Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989) User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35 (8), 982–1003.
- Decker, G., Overdick, H. & Weske, M. (2008) Oryx – An open modeling platform for the BPM community. In: Dumas, M., Reichert, M. & Shan, M.-C. (eds.) *Business Process Management*. Berlin / Heidelberg, Springer, pp. 382–385.
- Denning, P. J. (1997) A new social contract for research. *Communications of the ACM*, 40 (2), 132–134.
- Denscombe, M. (2010) *The Good Research Guide: For Small-Scale Social Research Projects*. New York, Open University Press.
- Dumas, M., La Rosa, M., Mendling, J. & Reijers, H. A. (2013) *Fundamentals of Business Process Management*. Berlin, Springer.
- Dumas, M., Aalst, W. M. v. d. & Hofstede, A. H. t. (2005) *Process-Aware Information Systems: Bridging People and Software Through Process Technology*. Hoboken, John Wiley & Sons Inc.
- Dunn, C., Cherrington, J. O. & Hollander, A. (2005) *Enterprise Information Systems: A Pattern-Based Approach*. Boston, MA, McGraw-Hill/Irwin.
- Edirisuriya, A. (2009) *Design Support for e-Commerce Information Systems using Goal, Business and Process Modelling*. Department of Computer and Systems Science, Stockholm, Stockholm University, pp. 187.
- Eid-Sabbagh, R.-H., Kunze, M. & Weske, M. (2012) An open process model library. *Business Process Management Workshops*. Berlin Heidelberg, Springer, pp. 26–38.
- Elhadad, M., Balaban, M. & Sturm, A. (2008) Effective business process outsourcing: The Prosero approach. *International Journal of Interoperability in Business Information Systems*, 3 (1), 8-31.

- Elias, M., Bider, I. & Johannesson, P. (2014) Using fractal process-asset model to design the process architecture of an enterprise: Experience report. In: Bider, I., Gaaloul, K., Krogstie, J., Nurcan, S., Proper, H., Schmidt, R., & Soffer, P. (eds.) *Enterprise, Business-Process and Information Systems Modeling*. Berlin Heidelberg, Springer, pp. 287–301.
- Elias, M. & Johannesson, P. (2013) A context-based process semantic annotation model for a process model repository. *Business Process Management Journal*, 19 (3), 404–430.
- Elias, M. & Johannesson, P. (2012a) An empirical assessment of the effect of context-based semantic annotation on process model discovery. In: Bajec, M. & Eder, J. (eds.) *Advanced Information Systems Engineering Workshops*. Berlin Heidelberg, Springer, pp. 366–382.
- Elias, M. & Johannesson, P. (2012b) A survey of process model reuse repositories. In: Dua, S., Gangopadhyay, A., Thulasiraman, P., Straccia, U., Shepherd, M., & Stein, B. (eds.) *Information Systems, Technology and Management. Grenoble, France*. Berlin Heidelberg, Springer, pp. 64–76.
- Elias, M. & Shahzad, K. (2010) Using multi-criteria decision making to choose process representation format for a process repository. In: Abramowicz, W., Tolksdorf, R. & Węcel, K. (eds.) *Business Information Systems Workshops*. Berlin Heidelberg, Springer, pp. 19–24.
- Elias, M., Shahzad, K. & Johannesson, P. (2010) A business process metadata model for a process model repository. In: Bider, I., Halpin, T., Krogstie, J., Nurcan, S., Proper, E., Schmidt, R., & Ukor, R. (eds.) *Enterprise, Business-Process and Information Systems Modeling*. Berlin Heidelberg, Springer, pp. 287–300.
- Embley, D. W. & Goldstein, R. C. (1997) *Proceedings of Conceptual Modeling-ER'97: 16th International Conference on Conceptual Modeling, Los Angeles, CA, USA*, Springer.
- Erl, T. (2005) *Service-Oriented Architecture (SOA): Concepts, Technology, and Design*. Boston, Prentice Hall.
- Ezran, M., Morisio, M. & Tully, C. (2002) *Practical Software Reuse*. London, Springer-Verlag.
- Fantinato, M., Toledo, M. B. F., Thom, L. H., Gimenes, I. M. S., Rocha, R. S., & Garcia, D. Z. G. (2012) A survey on reuse in the business process management domain. *International Journal of Business Process Integration and Management*, 6 (1), 52-76.
- Filipowska, A., Kaczmarek, M., Kowalkiewicz, M., Markovic, I. & Zhou, X. (2009) Organizational ontologies to support semantic business process management. In: *Proceedings of the 4th International Workshop on Semantic Business Process Management*. Heraklion, Greece, ACM, pp. 35–42.
- Flores, M., Flores, K., Cabello, A., Terzi, S. & Rossi, M. (2012) Understanding the approaches to create a process architecture for lean thinking. In: *Proceedings of the 18th International Conference on Engineering, Technology and Innovation*. IEEE, pp. 1–10.
- Forster, A. & Engels, G. (2003) Quality ensuring development of software processes. In: Oquendo, F. (ed.) *Software Process Technology*. Berlin Heidelberg, Springer, pp. 62–73.

- Fox, M. S., Barbuceanu, M. & Gruninger, M. (1996) An organisation ontology for enterprise modeling: Preliminary concepts for linking structure and behaviour. *Computers in Industry*, 29 (1), 123–134.
- Frolov, V., Mengel, D., Bandara, W., Sun, Y. & Ma, L. (2010) Building an ontology and process architecture for engineering asset management. In: Kiritsis, D., Emmanouilidis, C., Koronios, A. & Mathew, J. (eds.) *Engineering Asset Lifecycle Management*. London, Springer, pp. 86–97.
- Gao, S. & Krogstie, J. (2009) A combined framework for development of business process support systems. In: Persson, A. & Stirna, J. (eds.) *The Practice of Enterprise Modeling*. Berlin Heidelberg, Springer, pp. 115–129.
- Geerts, G. L. & McCarthy, W. E. (2000) *The Ontological Foundation of REA Enterprise Information Systems*. Technical Report, Michigan State University.
- Ghani, A., Azim, A., Koh, T. W., Muketha, G. M. & Wong, P. W. (2008) Complexity metrics for measuring the understandability and maintainability of business process models using goal-question-metric (GQM). *International Journal of Computer Science and Network Security*, 8 (5), 219–225.
- Give Team. (2010) *Insightmaker*. [Online] Available from: <http://insightmaker.com/>. [Accessed: 5th April 2014].
- Glushko, R. J. & McGrath, T. (2008) *Document engineering: analyzing and designing documents for business informatics and Web services*. Boston, The MIT Press.
- Gruber, T. R. (1995) Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, 43 (5), 907–928.
- Gruber, T. R. (1993) A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5 (2), 199–220.
- Gschwind, T., Koehler, J. & Wong, J. (2008) Applying patterns during business process modeling. In: Dumas, M., Reichert, M. & Shan, M.-C. (eds.) *Business Process Management*. Berlin Heidelberg, Springer, pp. 4–19.
- Guerin, S. (2005) Organizing business knowledge: The MIT process handbook. *Journal of Product Innovation Management*, 22 (2), 218–219.
- Hammer, M. & Stanton, S. (1999) How process enterprises really work. *Harvard Business Review*, 77, 108–120.
- Hammer, M. & Champy, J. (1993) *Reengineering the Corporation: A Manifesto for Business Revolution*. New York, HarperCollins.
- Havey, M. (2005) *Essential Business Process Modeling*. Cambridge, O'Reilly Media, Inc.
- Hepp, M. & Roman, D. (2007) An ontology framework for semantic business process management. In: *Proceedings of the 8th Wirtschaftsinformatik International Conference*. Karlsruhe, p. 423-440.
- Hepp, M., Leymann, F., Domingue, J., Wahler, A. & Fensel, D. (2005) Semantic business process management: A vision towards using semantic Web services for business process management. In: *Proceedings of IEEE International Conference on e-Business Engineering, (ICEBE 2005), Beijing, China*, pp. 535–540.
- Herrmann, C. & Kurz, M. (2011) Adaptive case management: Supporting knowledge intensive processes with IT systems. In: Schmidt, W. (ed.) S-

- BPM ONE – Learning by Doing – Doing by Learning*. Berlin Heidelberg, Springer, pp. 80–97.
- Hevner, A. R., March, S. T., Park, J. & Ram, S. (2004) Design science in information systems research. *MIS Quarterly*, 28 (1), 75–105.
- Hevner, A. R. & March, S. T. (2003) The information systems research cycle. *Computer*, 36 (11), 111–113.
- Hofreiter, B. & Huemer, C. (2006) From a UMM business process model to a business environment specific ebXML process. *Journal of E-Commerce Research*, 7 (3), 138–153.
- Holschke, O., Rake, J. & Levina, O. (2009) Granularity as a cognitive factor in the effectiveness of business process model reuse. In: Dayal, U., Eder, J., Koehler, J. & Reijers, H. (eds.) *Business Process Management*. Berlin / Heidelberg, Springer, pp. 245–260.
- Hornung, T., Koschmider, A. & Oberweis, A. (2009) A recommender system for business process models. In: *Proceedings of the 17th Annual Workshop on Information Technologies & Systems (WITS)*. Montreal, Canada, pp. 127–132
- Huang, A. H. (2003) Effects of multimedia on document browsing and navigation: An exploratory empirical investigation. *Information & Management*, 41 (2), 189–198.
- Huat Lim, S., Juster, N. & de Pennington, A. (1997) Enterprise modelling and integration: A taxonomy of seven key aspects. *Computers in Industry*, 34 (3), 339–359.
- IBM Corporation. (2004) *IBM Process Repository*. [Online] Available from: http://publib.boulder.ibm.com/infocenter/wchelp/v5r6m1/index.jsp?topic=/com.ibm.commerce.business_process.doc/concepts/processPrice_order.htm . [Accessed 10th September 2010.]
- IEEE. (2007) *Recommended Practice for Architectural Description of Software-intensive Systems*, ANSI/IEEE Std 1471-ISO/IEC 42010. Institute of Electrical and Electronics Engineers.
- Indulska, M., Green, P., Recker, J. & Rosemann, M. (2009) Business process Mmodeling: Perceived benefits. In: Laender, A., Castano, S., Dayal, U., Casati, F. & de Oliveira, J. (eds.) *Conceptual Modeling (ER 2009)*. Berlin / Heidelberg, Springer, pp. 458–471.
- ISO/IEC/(IEEE). (2011) ISO/IEC/IEEE 42010: Systems and software engineering – Architecture description. [Online] Available from: <http://www.iso-architecture.org/ieee-1471/> [Accessed 10th September 2014.]
- ITU-T, R. Z. & Recommendation, Z. (2000) 100: Specification and Description Language (SDL). In: *International Telecommunication Union*.
- Jablonski, S. & Bussler, C. (1996) *Workflow Management: Modeling Concepts, Architecture and Implementation*. London, UK, International Thomson Computer Press.
- Jayaweera, P. M. (2004) A Unified Framework for e-Commerce Systems Development: Business Process Pattern Perspective. *Department of Computers and Systems Sciences*. Stockhom: Sockholm University, pp. 185.
- Jeston, J. & Nelis, J. (2014) *Business Process Management: Practical Guide to Successful Implementation*. New York, Routledge.

- Jin, T., Wang, J., La Rosa, M., Ter Hofstede, A. & Wen, L. (2013) Efficient querying of large process model repositories. *Computers in Industry*, 64 (1), 41-49.
- Joahannesson, P. & Perjons, E. (2012) *A Design Science Primer*. North Charleston, SC, CreateSpace Independent Publishing Platform.
- Joahannesson, P. & Perjons, E. (2001) Design principles for process modelling in enterprise application integration. *Information Systems*, 26 (3), 165–184.
- Johansson, H. J., Machugh, P., Pendlebury, A. J. & Wheeler, W. A. (1993) *Business Process Reengineering: Breakpoint Strategies for Market Dominance*. Chichester, Wiley.
- Minrong, S., John, M. & Ismailcem, A. (2001) RepoX: An XML Repository for Workflow Design and Specifications. Technical Report #UGA-CS-LSDIS-TR-01- 011, University of Georgia, August 2001.
- Jussi, V., Jana, K. & Frank, L. (2006) Repository for business processes and arbitrary associated metadata. In: *Proceedings of the 4th International Conference on Business Process Management*. Vienna, Austria, pp. 25–31.
- Jussi, V. (2004) Building and querying a repository of bpm process specifications. *Institute Eurecom*. Helsinki University of Technology, and University of Nice-Sophia Antipolis.
- Jung, J.-Y., Bae, J. & Liu, L. (2009) Hierarchical clustering of business process models. *International Journal of Innovative Computing, Information and Control*, 5 (12), 1349-4198.
- Kandfer, M. & Shapiro, S. (2009) An F-measure for context-based information retrieval. *International Symposium on Logical Formalization of Commonsense Reasoning*. Toronto, CA, pp. 79–84.
- Kärger, P., Ullrich, C. & Melis, E. (2006) Integrating learning object repositories using a mediator architecture. In: Nejdil, W. & Tochtermann, K. (eds.) *Innovative Approaches for Learning and Knowledge Sharing*. Berlin Heidelberg, Springer, pp. 185–197.
- Khanna, A. (2013) Managing Unpredictability using BPM for Adaptive Case Management. [Online] Available from: <http://www.oracle.com/us/technologies/bpm/bpm-for-adaptive-case-mgmt-1972799.pdf> [Accessed 10th September 2014.]
- Kitchenham, B. & Charters, S. (2007) Guidelines for Performing Systematic Literature Reviews in Software Engineering. School of Computer Science and Mathematics, Keele University. Technical Report ESBE-2007-1.
- Kitchenham, B. (2004) Procedures for performing systematic reviews. *Technical Report TR/SE-0401 and NICTA Technical Report 0400011T.1*. Keele, UK: Keele University and National ICT Australia Ltd, pp. 33.
- Ko, R. K. L. (2009) A computer scientist's introductory guide to business process management (BPM). *Crossroads*, 15 (4), 11–18.
- Korherr, B. (2008) Business process modelling – Languages, goals, and variabilities. *Faculty of Informatics*. Vienna, Vienna University of Technology, pp. 115.
- Kotonya, G. & Sommerville, I. (1998) *Requirement Engineering: Processes and Techniques*. New York, Wiley.
- Krogstie, J. & Sjølvberg, A. (2003) *Information systems engineering: Conceptual modeling in a quality perspective*. Kompendiumforlaget, Trondheim, Norway.

- Krogstie, J. (2012) *Model-based development and evolution of information systems: A Quality Approach*. Berlin Heidelberg, Springer.
- Krueger, C. W. (1992) Software reuse. *ACM Computing Surveys*, 24 (2), 131–183.
- Kueng, P. & Kawalek, P. (1997) Goal-based business process models: Creation and evaluation. *Business Process Management Journal*, 3 (1), 17–38.
- Kueng, P., Bichler, P., Kawalek, P. & Schrefl, M. (1996) How to compose an object-oriented business process model? In: Brinkkemper, S., Lyytinen, K. & Welke, R. (eds.) *Method Engineering: Principle of method construction and tool support*. London, Chapman & Hall, pp. 94–110.
- Kunze, M. & Weske, M. (2011) Metric Trees for Efficient Similarity Search in Large Process Model Repositories. In: zur Muehlen, M. & Su, J. (eds) *Business Process Management Workshops*. Berlin Heidelberg, Springer, pp. 535-546.
- Kurniawan, T., Ghose, A., L ê, L.-S., & Dam, H. (2012) On formalizing inter-process relationships. In: Daniel, F., Barkaoui, K. & Dustdar, S. (eds.) *Business Process Management Workshops*. Berlin Heidelberg, Springer, pp. 75–86.
- La Rosa, M., Reijers, H. A., van der Aalst, W. M. P., Dijkman, R. M., Mendling, J., Dumas, M. & Garcia-Banuelos, L. (2011a) APROMORE: An advanced process model repository. *Expert Systems with Applications*, 38 (6), 7029–7040.
- La Rosa, M., ter Hofstede, A. H. M., Wohed, P., Reijers, H. A., Mendling, J., & van der Aalst, W. M. P. (2011b) Managing process model complexity via concrete syntax modifications. *IEEE Transactions on Industrial Informatics*, 7 (2), 255–265.
- La Rosa, M. (2009) Managing Variability in Process-Aware Information Systems. *Faculty of Science and Technology*. Brisbane: Queensland University of Technology, pp. 164.
- Lakhotia, A. (1993) Understanding someone else's code: Analysis of experiences. *Journal of Systems and Software*, 23 (3), 269–275.
- Laue, R. & Mendling, J. (2010) Structuredness and its significance for correctness of process models. *Information Systems and E-Business Management*, 8 (3), 287-307.
- Lautenbacher, F. & Bauer, B. (2007) A Survey on Workflow Annotation & Composition Approaches. In: *Proceedings of the Workshop on Semantic Business Process and Product Lifecycle Management (SBPM 2007)*, Innsbruck, Austria, pp. 12-23.
- Lawrence, P. (1997) *Workflow Handbook 1997*. New York, John Wiley & Sons, Inc.
- Leon, A. (2012) *Enterprise Resource Planning*. New York, McGraw-Hill Education.
- Liao, L. & Leung, H. K. (2007) An ontology-based business process modeling methodology. In: *Proceedings of the 3rd International Association of Science and Technology for Development (IASTED 2007) on Advances in Computer Science and Technology (ACST 2007)*. Phuket, Thailand.
- Liao, Y., Lezoche, M., Panetto, H. & Boudjlida, N. (2011) Semantic annotation model definition for systems interoperability. In: Meersman, R., Dillon, T. & Herrero, P. (eds.) *On the Move to Meaningful Internet Systems: OTM 2011 Workshops*. Berlin Heidelberg, Springer, pp. 61–70.

- Lin, Y. (2008) Semantic annotation for process models: Facilitating process knowledge management via semantic interoperability. *Department of Computer and Information Science*. Trondheim: NTNU, Norway, pp. 239.
- Lin, Y. & Sølvsberg, A. (2007) Goal annotation of process models for semantic enrichment of process knowledge. In: Krogstie, J., Opdahl, A. & Sindre, G. (eds.) *Advanced Information Systems Engineering*. Berlin / Heidelberg, Springer, pp. 355–369.
- Lin, Y. & Strasunskas, D. (2005) Ontology-based semantic annotation of process templates for reuse. In: *Proceedings of 10th International Workshop on Exploring Modeling Methods for Systems Analysis and Design (EMMSAD)*, Porta, Portugal, pp. 31 - 37.
- Lindland, O. I., Sindre, G. & Solvberg, A. (1994) Understanding quality in conceptual modeling. *Software, IEEE*, 11 (2), 42–49.
- List, B. & Korherr, B. (2006) An evaluation of conceptual business process modelling languages. In: *Proceedings of the 2006 ACM symposium on Applied computing*. Dijon, France: ACM, pp. 1532–1539.
- Louvel, J., Templier, T. & Boileau, T. (2012) *Restlet in Action*. New York, Manning Publications.
- Lu, R. & Sadiq, S. (2007) A survey of comparative business process modeling approaches. In: Abramowicz, W. (ed.) *Business Information Systems*. Berlin / Heidelberg, Springer, pp. 82–94.
- Lucena, M., Santos, E., Silva, C., Alencar, F., Silva, M. J., & Castro, J. (2008) Towards a unified metamodel for i*. In: *Proceedings of Research Challenges in Information Science (RCIS)*, pp. 237–246.
- Lucrédio, D., Brito, K. S., Alvaro, A., Garcia, V. C., Almeida, E. S., Fortes, R. P. M., and ´ Meira, S. R. L. (2008) Software reuse: The Brazilian industry scenario. *Journal of Systems and Software*, 81 (6), 996–1013.
- Ma, Z. & Leymann, F. (2008) A lifecycle model for using process fragment in business process modeling. In: *Proceedings of the 9th International Workshop on Business Process Modeling, Development and Support (BPMDS'08), held in conjunction with CAiSE'08*. Montpellier, France, pp. 1-9.
- Ma, Z., Wetzstein, B., Anicic, D., Heymans, S. & Leymann, F. (2007) Semantic business process repository. In: *CEUR Workshop Proceedings*, pp. 92-100.
- McCarthy, W. E. (1982) The REA accounting model: A generalized framework for accounting systems in a shared data environment. *The Accounting Review*, 57,554–578.
- Malone, T. W., Kevin, C. & George, A. H. (2003) *Organizing Business Knowledge: The MIT Process Handbook*, Cambridge, MA, MIT Press.
- Malone, T. W., Crowston, K., Lee, J., Pentland, B., Dellarocas, C., Wyner, G., Quimby, J., Osborn, C. S., Bernstein, A., Herman, G., Klein, M., & O'Donnell, E. (1999) Tools for inventing organizations: Toward a handbook of organizational processes. *Management Science*, 45 (3), 425–443.
- March, S. T. & Smith, G. F. (1995) Design and natural science research on information technology. *Decision Support Systems*, 15 (4), 251–266.
- Markov, I. & Kowalkiewicz, M. (2008) Linking Business Goals to Process Models in Semantic Business Process Modeling. In: *Proceedings of the 12th*

- International IEEE Enterprise Distributed Object Computing Conference*. Munich, Germany, IEEE Computer Society, pp. 332-338.
- Markovic, I. & Pereira, A. (2008a) Towards a formal framework for reuse in business process modeling. In: ter Hofstede, A., Benatallah, B. & Paik, H.-Y. (eds.) *Business Process Management Workshops*. Berlin Heidelberg, Springer, pp. 484-495.
- Markovic, I. & Pereira, A. C. (2008b) Towards a formal framework for reuse in business process modeling. In: *Proceedings of the 2007 International Conference on Business Process Management*. Brisbane, Australia, Springer-Verlag, pp. 484-495.
- Measuring Usability LLC. (2004) *ANAOVA Calculator*. [Online] Available from: <http://www.usablestats.com/calcs/2samplet>. [Accessed 10th December, 2014].
- Melcher, J., Mendling, J., Reijers, H. A. & Seese, D. (2010) On measuring the understandability of process models. In: Rinderle-Ma, S., Sadiq, S. & Leymann, F. (eds.) *Business Process Management Workshops*. Berlin Heidelberg, Springer, pp. 465-476.
- Mendling, J., Strembeck, M. & Recker, J. (2012) Factors of process model comprehension: Findings from a series of experiments. *Decision Support Systems*, 53 (1), 195-206.
- Mendling, J. & Strembeck, M. (2008) Influence factors of understanding business process models. In: Abramowicz, W. & Fensel, D. (eds.) *Business Information Systems*. Berlin Heidelberg, Springer, pp. 142-153.
- Mendling, J. & Nüttgens, M. (2006) EPC markup language (EPML): An XML-based interchange format for event-driven process chains (EPC). *Information Systems and E-Business Management*, 4 (3), 245-263.
- MIT Process Handbook Project. (2001) *MIT Process Handbook*. [Online] Available from: <http://process.mit.edu/Default.asp>. [Accessed 7th September 2010].
- Moher, T. G., Mak, D., Blumenthal, B. & Levanthal, L. (1993) Comparing the comprehensibility of textual and graphical programs. The Case of Petri Nets. In: *Proceedings of fifth Workshop on Empirical Studies of Programmers*. Norwood, NJ, Ablex, pp. 137-161.
- Mondary, T., Bouffier, A. & Nazarenko, A. (2007) Between browsing and search, a new model for navigating through large documents. In: *Proceedings of the 2nd European Cognitive Science Conference*. Greece, pp. 634-639.
- Moody, D. L. (2003) The method evaluation model: A theoretical model for validating information systems design methods. In: *Proceedings of the 11th European Conference on Information Systems*. Naples, Italy, pp. 1-17.
- Moody, D. L., Sindre, G., Brasethvik, T. & Solvberg, A. (2003) Evaluating the quality of information models: Empirical testing of a conceptual model quality framework. In: *Proceedings of the 25th International Conference on Software Engineering (ICSE'03)*. IEEE Computer Society, pp. 295-305.
- Morrow, R. A. & Brown, D. D. (1994) *Contemporary Social Theory: Critical Theory and Methodology*. Thousand Oaks, CA, Sage Publications.
- Nelson, K. & Nelson, H. (2003) The need for a strategic ontology. In: *Proceedings of the MIS (Management Information Systems) Quarterly Special Issue Workshop on Standard Making: A Critical Research Frontier for Information Systems*. pp. 12-14.

- Netjes, M., Reijers, H. A. & van der Aalst, W. M. (2006) Supporting the BPM life-cycle with FileNet. *Proceedings of the CAiSE*. Citeseer, pp. 497–508.
- Noy, N. F. & McGuinness, D. L. (2001) Ontology development 101: A guide to creating your first ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880.
- Object Management Group. (2006) *Business Process Modeling Notation*. [Online] Available from: <http://www.bpmn.org/>. [Accessed 10th September 2010].
- Object Management Group. (2004) *UML 2.0 Superstructure Specification*. [Online] Available from: <http://omg.org/cgi-bin/doc?ptc/2004-10-02>. [Accessed 7th September 2010].
- Oei12, J., van Hemmen, L., Falkenberg, E. D., & Brinkkemper, S. (1992) The meta model hierarchy: A framework for information systems concepts and techniques. Nijmegen, The Netherlands: Department of Informatics, Faculty of Mathematics and Informatics, University of Nijmegen.
- Oryx. (2008) *The Oryx Project*. [Online] Available from: <http://bpt.hpi.unipotsdam.de/Oryx/WebHome>. [Accessed 13th September 2010].
- Ottensooser, A., Fekete, A., Reijers, H. A., Mendling, J. & Menictas, C. (2012) Making sense of business process descriptions: An experimental comparison of graphical and textual notations. *Journal of Systems and Software*, 85 (3), 596–606.
- Panagacos, T. (2012) *The Ultimate Guide to Business Process Management: Everything You Need to Know and How to Apply it to Your Organization*. Washington, CreateSpace Independent Publishing Platform.
- Pedrinaci, C., Domingue, J. & Alves de Medeiros, A. (2008) A core ontology for business process analysis. In: Bechhofer, S., Hauswirth, M., Hoffmann, J. r. & Koubarakis, M. (eds.) *The Semantic Web: Research and Applications*. Berlin Heidelberg, Springer, pp. 49–64.
- Peffers, K., Rothenberger, M., Tuunanen, T. & Vaezi, R. (2012) Design science research evaluation. In: Peffers, K., Rothenberger, M. & Kuechler, B. (eds.) *Design Science Research in Information Systems. Advances in Theory and Practice*. Berlin Heidelberg, Springer, pp. 398–410.
- Peffers, K., Tuunanen, T., Rothenberger, M. & Chatterjee, S. (2007) A design science research methodology for information systems research. In: *Journal of Management Information System*, 24 (3), 45–77.
- Perjons, E. (2011) *ModelDriven Process Design: Aligning Value Networks, Enterprise Goals, Services and IT Systems*. Stockholm, Department of Computer and Systems Sciences, Stockholm University.
- Peters, N. (2007) *Oryx stencil set specification*. Potsdam, Germany.
- Petre, M. (1995) Why looking isn't always seeing: Readership skills and graphical programming. *Communications of the ACM*, 38 (6), 33–44.
- Petri, C. A. (1962) *Kommunikation mit automaten*. University of Born, Germany.
- Poels, G., Maes, A., Gailly, F. & Paemeleire, R. (2005) Measuring the perceived semantic quality of information models. In: Akoka, J., Liddle, S., Song, I., Bertolotto, M., Comyn-Wattiau, I., Heuvel, W., Kolp, M., Trujillo, J., Kop, C., & Mayr, H. eds.) *Perspectives in Conceptual Modeling*. Berlin Heidelberg, Springer, pp. 376–385.

- Polyvyanyy, A., Smirnov, S. & Weske, M. (2010) Business process model abstraction. In: Brocke, J. & Rosemann, M. (eds.) *Handbook on Business Process Management I*. Berlin Heidelberg, Springer, pp. 149–166.
- Polyvyanyy, A., Smirnov, S. & Weske, M. (2009) The triconnected abstraction of process models. In: Dayal, U., Eder, J., Koehler, J. & Reijers, H. (eds.) *Business Process Management*. Berlin Heidelberg, Springer, pp. 229–244.
- Polyvyanyy, A., Smirnov, S. & Weske, M. (2008) Process model abstraction: A slider approach. *Proceeding of the 12th Int'l IEEE Enterprise Distributed Object Computing Conference (EDOC '08)*. München, Germany: IEEE, pp. 325–331.
- Porter, M. (2008) *On Competition*. Boston, MA, Harvard Business School Press.
- Porter, M. (1985) *Competitive Advantage*. New York: Free Press.
- Prieto-Diaz, R. & Freeman, P. (1987) Classifying software for reusability. *Software, IEEE*, 4 (1), 6–16.
- Purchase, H. (1997) Which aesthetic has the greatest effect on human understanding? In: DiBattista, G. (ed.) *Graph Drawing*. Berlin Heidelberg, Springer, pp. 248–261.
- Qiao, M., Akkiraju, R. & Rembert, A. (2011) Towards Efficient Business Process Clustering and Retrieval: Combining Language Modeling and Structure Matching. In: Rinderle-Ma, S., Toumani, F. & Wolf, K. (eds) *Business Process Management*. Berlin Heidelberg, Springer, pp. 199–214.
- Rademakers, T. (2012) *Activiti in Action: Executable business processes in BPMN 2.0*. New York, Manning Publications.
- Recker, J. C., Rosemann, M., Indulska, M. & Green, P. (2009) Business process modeling: A comparative analysis. *Journal of the Association for Information Systems*, 10 (4), 333–363.
- Reijers, H. A. & Mendling, J. (2011) A study into the factors that influence the understandability of business process models. *IEEE Transactions on Systems, Man and Cybernetics – Part A*, 41 (3), 449–462.
- Reijers, H. A., Freytag, T., Mendling, J. & Eckleder, A. (2011) Syntax highlighting in business process models. *Decision Support Systems*, 51 (3), 339–349.
- Reijers, H. A. (2006) Implementing BPM systems: The role of process orientation. *Business Process Management Journal*, 12 (4), 389–409.
- Rescher, N. (1997) *Methodological Pragmatism: Systems-Theoretic Approach to the Theory of Knowledge*. Oxford, Blackwell Publishers.
- Riehle, D. & Zllighoven, H. (1996) Understanding and using patterns in software development. *Theory and Practice of Object System*, 2 (1), 3–13.
- Ritchie, J. & Spencer, L. (2002) Qualitative data analysis for applied policy research. *The Qualitative Researcher's Companion*, pp. 305–329.
- Rodrigues, J. A., Souza, J. M., Zimbrão, G., Xexeo, G., Neves, E., & Pinheiro, W. A. (2006) A P2P Approach for business process modelling and reuse. In: Eder, J. & Dustdar, S. (eds.) *Business Process Management Workshops*. Berlin / Heidelberg, Springer, pp. 297–307.
- Rummler, G. A. and Brache, A. P. (1995) *Improving Performance: How to Manage the White Space on the Organizational Chart*, San Francisco, CA: Jossey-Bass.
- SAP AG. (2007) *SAP Business Map*. [Online] Available from: http://help.sap.com/saphelp_sm40/helpdata/EN/5e/c8145e3a9d9340913099159d80fc87/frameset.htm. [Accessed 7th September 2010].

- Sarshar, K. & Loos, P. (2005) Comparing the control-flow of EPC and petri net from the end-user perspective. In: van der Aalst, W. P., Benatallah, B., Casati, F. & Curbera, F. (eds.) *Business Process Management*. Berlin Heidelberg, Springer, pp. 434–439.
- Scheer, A.-W. (2000) *ARIS : Business Process Modeling*. Berlin, Alemania Springer.
- Schon, D. (1984) *The Reflective Practitioner: How Professionals Think In Action*. New York, Basic Books.
- Schrepfer, M., Wolf, J., Mendling, J. & Reijers, H. (2009) The impact of Secondary Notation on Process Model understanding. In: Persson, A. & Stirna, J. (eds.) *The Practice of Enterprise Modeling*. Berlin Heidelberg, , pp. 161–175.
- Schumm, D., Karastoyanova, D., Kopp, O., Leymann, F., Sonntag, M., & Strauch, S. (2011a) Process fragment libraries for easier and faster development of process-based applications. *Journal of Systems Integration*, 2 (1), 39–55.
- Schumm, D., Karastoyanova, D., Leymann, F. & Strauch, S. (2011b) Fragmento: Advanced process fragment library. In: Pokorny, J., Repa, V., Richta, K., Wojtkowski, W., Linger, H., Barry, C., & Lang, M (eds.) *Information Systems Development*. New York, Springer, pp. 659–670.
- Shahzad, K., Elias, M. & Johannesson, P. (2010) Requirements for a business process model repository: A stakeholders' perspective. In: Abramowicz, W. & Tolksdorf, R. (eds.) *Business Information Systems*. Berlin Heidelberg, Springer, pp. 158–170.
- Shahzad, K., Andersson, B., Bergholtz, M., Edirisuriya, A., Ilayperuma, T., Jayaweera, P., & Johannesson, P. (2009) Elicitation of requirements for a business process model repository. *Business Process Management Workshops*. Berlin Heidelberg, Springer, pp. 44–55.
- Shiva, S. G. & Shala, L. A. (2008) Using semantic Wikis to support software reuse. *Journal of Software*, 3 (4),1-8.
- Simon, H. (1981) *The Sciences of the Artificial*. Cambridge, MA: The MIT Press.
- Smirnov, S., Reijers, H., Weske, M. & Nugteren, T. (2012) Business process model abstraction: A definition, catalog, and survey. *Distributed and Parallel Databases*, 30 (1),63–99.
- Smirnov, S., Reijers, H. & Weske, M. (2011) A semantic approach for business process model abstraction. In: Mouratidis, H. & Rolland, C. (eds.) *Advanced Information Systems Engineering*. Berlin Heidelberg, Springer, pp. 497–511.
- Smirnov, S. (2010) *Business Process Model Abstraction: Theory and Practice*. Potsdam, University of Potsdam.
- Stenning, K. & Oberlander, J. (1997) A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, 19 (1), 97–140.
- Wegner, B. & Monk, E. (2008) *Enterprise Resource Planning*. Boston, Cengage Learning.
- SUPER. (2004) *Semantics Utilized for process management with and between enterprises* [Online] Available from: <http://www.ip-super.org/>. [Accessed 7th September 2010].
- SUPER. (2007) *Semantics utilized for process management within and between enterprises*. [Online] Available from: <http://www.ip-super.org/>. [Accessed 27th May 2011].

- Supply Chain Council. (2003) *The Phios Process Repository*. [Online] Available from: <http://repository.phios.com/SCOR/>. [Accessed 7th September 2010.]
- Swenson, K. D., Palmer, N. & Pucher, M. J. (2012) *How Knowledge Workers Get Things Done: Real-World Adaptive Case Management*. Lighthouse Point, FL, Future Strategies Inc.
- Team, A. (2010) *Activiti BPM Platform*. [Online] Available from: <http://www.activiti.org/>. [Accessed 21st July 21 2014].
- Tekinerdogan, B. & Aksit, M. (2002) Synthesis-based software architecture design. In: Aksit, M. (ed.) *Software Architectures and Component Technology*. US, Springer, pp. 143–173.
- Thom, L. H., Reichert, M. & Iochpe, C. (2009) Activity patterns in process-aware information systems: Basic concepts and empirical evidence. *International Journal of Business Process Integration and Management*, 4 (2), 93–110.
- Thomas, O. & Fellmann, M. (2006) Semantic event-driven process chains. In: *Proceedings of the Workshop on Semantic Business Process and Product Lifecycle Management (SBPM 2006) in conjunction with 3rd European Semantic Web Conference (ESWC 2006), Budva, Montenegro*, pp. 64–75.
- TM Forum. (2009) *Enhanced telecom operations model (eTOM) process classification scheme*. [Online] Available from: <http://www.tmforum.org/IntegrationFramework/6637/home.html>. [Accessed 7th September 2014].
- Tudorache, T., Nyulas, C., F. Noy, N. & Musen, M. A. (2011) WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the Web. *Semantic web 4*. California, USA.
- UN/CEFACT. (2001) *Context and re-usability of core components, Version 1.04* [Online] Available from: <http://www.ebxml.org/specs/ebCNTXT.pdf> 1. [Accessed 13th September 2010].
- UN/CEFACT. (2003) *UN/CEFACT Modeling Methodology (UMM) User Guide*. [Online] Available from: http://www.unece.org/fileadmin/DAM/cefact/umm/UMM_userguide_220606.pdf. [Accessed 7th September 2014].
- UN/CEFACT. (2005) *Common Business Process Catalog (CBPC). Technical Specification, Version 1.0*. [Online] Available from: http://www.uncefactforum.org/TBG/TBG14/TBG14Documents/cbpc-technical-specification-v1_0-300905-11.pdf. [Accessed 7th September 2014].
- UN/CEFACT. (2011) *UN/CEFACT Modelling Methodology (UMM)*. [Online] Available from: http://www.unece.org/cefact/umm/umm_index.html. [Accessed 12th May, 2014 2014].
- van der Aalst, W. M. (2013) Business process management: A comprehensive survey. *ISRN Software Engineering*, doi:10.1155/2013/507984, pp. 1–37.
- van der Aalst, W. M. P. (2011) On the Representational Bias in Process Mining. In: *Proceedings of 20th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE). Paris, France*. IEEE Computer Society, pp. 2–7.
- van der Aalst, W. & Stahl, C. (2011) *Modeling Business Processes: A Petri Net-Oriented Approach*. Cambridge, MA, MIT Press.
- van der Aalst, W. M. P. & ter Hofstede, A. H. M. (2005) YAWL: Yet another workflow language. *Information Systems*, 30 (4), 245–275.

- van Der Aalst, W. & Van Hee, K. M. (2004) *Workflow management: models, methods, and systems*, Cambridge, MA: MIT Press.
- van der Aalst, W. M. P., ter Hofstede, A. H. M., Kiepuszewski, B. & Barros, A. P. (2003a) Workflow patterns. *Distributed and Parallel Databases*, 14 (1), 5–51.
- van der Aalst, W. P., ter Hofstede, A. M. & Weske, M. (2003b) Business process management: A survey. In: van der Aalst, W. P. & Weske, M. (eds.) *Business Process Management*. Berlin Heidelberg, Springer, pp. 1–12.
- van der Aalst, W. M. P. & ter Hofstede, A. H. M. (1999) *Workflow Patterns Initiative*. [Online] Available from: <http://www.workflowpatterns.com>. [Accessed 2nd September, 2014.]
- Vom Brocke, J. & Rosemann, M. (2010) *Handbook on business process management 1: Introduction, Methods, and Information Systems (International Handbooks on Information Systems)*, Berlin, Springer.
- Vulcu, G., Bhiri, S., Derguech, W. & Ibáñez, M. J. (2011) Semantically-enabled business process models discovery. *International Journal of Business Process Integration and Management*, 5 (3), 257-272.
- Warner, J. (2000) In the catalogue ye go for men: evaluation criteria for information retrieval systems. *Aslib Proceedings*, 52 (2), 76–82.
- Weidlich, M., Smirnov, S., Wiggert, C. & Weske, M. (2011) Flexab-flexible business process model abstraction. *CAiSE Forum*. pp. 17–24.
- Weske, M. (2012) *Business Process Management: Concepts, Languages, Architectures*. Berlin, Springer.
- Weske, M. (2007) *Business Process Management: Concepts, Languages, Architectures*. Berlin, Springer.
- Wetzstein, B., Ma, Z., Filipowska, A., Kaczmarek, M., Bhiri, S., Losada, S., Lopez-Cob, J.-M., & Cicurel, L. (2007) Semantic business process management: A lifecycle based requirements analysis. In: *CEUR Workshop Proceedings*, pp. 1-11
- Wiederhold, G. (1992) Mediators in the architecture of future information systems. *Computer*, 25 (3), 38–49.
- Wikipedia Contributors. (2014a) Annotation. *Wikipedia – The Free Encyclopedia*.
- Wikipedia Contributors. (2014b) Business process. *Wikipedia –The Free Encyclopedia*.
- Wohead, P., Zdravkovic, J. & Johannesson, P. (2007) A Universal Repository for Process Models. Stockholm: Department of Computers and Systems Science, pp. 9.
- Wu, S. L., Xu, L. & He, W. (2009) Industry-oriented enterprise resource planning. *Enterprise Information Systems*, 3 (4),409–424.
- Yan, Z. (2012) Business process model repositories – Efficient process retrieval. Eindhoven, Eindhoven University of Technology, pp. 183.
- Yan, Z., Dijkman, R. & Grefen, P. (2010) Fast business process similarity search with feature-based similarity estimation. In: Meersman, R., Dillon, T. & Herrero, P. (eds.) *On the Move to Meaningful Internet Systems: OTM 2010*. Berlin Heidelberg, Springer, pp. 60–77.
- Yan, Z., Dijkman, R. & Grefen, P. (2012) Business process model repositories – Framework and survey. *Information and Software Technology*, 54 (4),380–395.

- Zdravkovic, J. (2009) *2nd IFIP WG 8.1 Working Conference on The Practice of Enterprise Modeling* [Online] Available from: <http://poem.dsv.su.se>. [Accessed 2nd September, 2014.]
- Zdravkovic, J. (2006) *Process Integration for the Extended Enterprise*. Stockholm: Royal Institute of Technology.
- Zhao, X. & Liu, C. (2007) Version management in the business process change context. *Proceedings Of The 5th International conference On Business Process Management*. Brisbane, Australia, Springer-Verlag, pp. 198-213.
- Zobel, J. (1998) How reliable are the results of large-scale information retrieval experiments?. In: *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. Melbourne, Australia, ACM, pp. 307–314.
- Zowghi, D. & Coulin, C. (2005) Requirements Elicitation: A Survey of Techniques, Approaches, and Tools. In: Aurum, A. k. & Wohlin, C. (eds) *Engineering and Managing Software Requirements*. Springer Berlin Heidelberg, pp. 19-46

DEPARTMENT OF COMPUTER AND SYSTEMS SCIENCES

Stockholm University/KTH

www.dsv.su.se

Ph.D. theses:

No 91-004 **Olsson, Jan**

An Architecture for Diagnostic Reasoning Based on Causal Models

No 93-008 **Orci, Terttu**

Temporal Reasoning and Data Bases

No 93-009 **Eriksson, Lars-Henrik**

Finitary Partial Definitions and General Logic

No 93-010 **Johannesson, Paul**

Schema Integration Schema Translation, and Interoperability in Federated Information Systems

No 93-018 **Wangler, Benkt**

Contributions to Functional Requirements Modelling

No 93-019 **Boman, Magnus**

A Logical Specification for Federated Information Systems

No 93-024 **Rayner, Manny**

Abductive Equivalential Translation and its Application to Natural-Language Database Interfacing

No 93-025 **Idestam-Almquist, Peter**

Generalization of Clauses

No 93-026 **Aronsson, Martin**

GCLA: The Design, Use, and Implementation of a Program Development

No 93-029 **Boström, Henrik**

Explanation-Based Transformation of Logic programs

No 94-001 **Samuelsson, Christer**

Fast Natural Language Parsing Using Explanation-Based Learning

No 94-003 **Ekenberg, Love**

Decision Support in Numerically Imprecise Domains

No 94-004 **Kowalski, Stewart**

IT Insecurity: A Multi-disciplinary Inquiry

No 94-007 **Asker, Lars**

Partial Explanations as a Basis for Learning

No 94-009 **Kjellin, Harald**

A Method for Acquiring and Refining Knowledge in Weak Theory Domains

No 94-011 **Britts, Stefan**

Object Database Design

No 94-014 **Kilander, Fredrik**

Incremental Conceptual Clustering in an On-Line Application

No 95-019 **Song, Wei**

Schema Integration: - Principles, Methods and Applications

No 95-050 **Johansson, Anna-Lena**

Logic Program Synthesis Using Schema Instantiation in an Interactive Environment

No 95-054 **Stensmo, Magnus**

Adaptive Automated Diagnosis

No 96-004 **Wærn, Annika**

Recognising Human Plans: Issues for Plan Recognition in Human - Computer Interaction

No 96-006 **Orsvärn, Klas**
Knowledge Modelling with Libraries of Task Decomposition Methods

No 96-008 **Dalianis, Hercules**
Concise Natural Language Generation from Formal Specifications

No 96-009 **Holm, Peter**
On the Design and Usage of Information Technology and the Structuring of Communication and Work

No 96-018 **Höök, Kristina**
A Glass Box Approach to Adaptive Hypermedia

No 96-021 **Yngström, Louise**
A Systemic-Holistic Approach to Academic Programmes in IT Security

No 97-005 **Wohed, Rolf**
A Language for Enterprise and Information System Modelling

No 97-008 **Gambäck, Björn**
Processing Swedish Sentences: A Unification-Based Grammar and Some Applications

No 97-010 **Kapidzic Cicovic, Nada**
Extended Certificate Management System: Design and Protocols

No 97-011 **Danielson, Mats**
Computational Decision Analysis

No 97-012 **Wijkman, Pierre**
Contributions to Evolutionary Computation

No 97-017 **Zhang, Ying**
Multi-Temporal Database Management with a Visual Query Interface

No 98-001 **Essler, Ulf**
Analyzing Groupware Adoption: A Framework and Three Case Studies in Lotus Notes Deployment

No 98-008 **Koistinen, Jari**
Contributions in Distributed Object Systems Engineering

No 99-009 **Hakkarainen, Sari**
Dynamic Aspects and Semantic Enrichment in Schema Comparison

No 99-015 **Magnusson, Christer**
Hedging Shareholder Value in an IT dependent Business society - the Framework BRITS

No 00-004 **Verhagen, Henricus**
Norm Autonomous Agents

No 00-006 **Wohed, Petia**
Schema Quality, Schema Enrichment, and Reuse in Information Systems Analysis

No 01-001 **Hökenhammar, Peter**
Integrerad Beställningsprocess vid Datasystemutveckling

No 01-008 **von Schéele, Fabian**
Controlling Time and Communication in Service Economy

No 01-015 **Kajko-Mattsson, Mira**
Corrective Maintenance Maturity Model: Problem Management

No 01-019 **Stirna, Janis**
The Influence of Intentional and Situational Factors on Enterprise Modelling Tool Acquisition in Organisations

No 01-020 **Persson, Anne**

Enterprise Modelling in Practice: Situational Factors and their Influence on Adopting a Participative Approach
 No 02-003 **Sneiders, Eriks**
 Automated Question Answering: Template-Based Approach
 No 02-005 **Eineborg, Martin**
 Inductive Logic Programming for Part-of-Speech Tagging
 No 02-006 **Bider, Ilia**
 State-Oriented Business Process Modelling: Principles, Theory and Practice
 No 02-007 **Malmberg, Åke**
 Notations Supporting Knowledge Acquisition from Multiple Sources
 No 02-012 **Männikkö-Barbutiu, Sirkku**
 SENIOR CYBORGS- About Appropriation of Personal Computers Among Some Swedish Elderly People
 No 02-028 **Brash, Danny**
 Reuse in Information Systems Development: A Qualitative Inquiry
 No 03-001 **Svensson, Martin**
 Designing, Defining and Evaluating Social Navigation
 No 03-002 **Espinoza, Fredrik**
 Individual Service Provisioning
 No 03-004 **Eriksson-Granskog, Agneta**
 General Metarules for Interactive Modular Construction of Natural Deduction Proofs
 No 03-005 **De Zoysa, T. Nandika Kasun**
 A Model of Security Architecture for Multi-Party Transactions
 No 03-008 **Tholander, Jakob**
 Constructing to Learn, Learning to Construct - Studies on Computational Tools for Learning
 No 03-009 **Karlgren, Klas**
 Mastering the Use of Gobbledygook - Studies on the Development of Expertise Through Exposure to Experienced Practitioners' Deliberation on Authentic Problems
 No 03-014 **Kjellman, Arne**
 Constructive Systems Science - The Only Remaining Alternative?
 No 03-015 **Rydberg Fåhræus, Eva**
 A Triple Helix of Learning Processes - How to cultivate learning, communication and collaboration among distance-education learners
 No 03-016 **Zemke, Stefan**
 Data Mining for Prediction - Financial Series Case
 No 04-002 **Hulth, Anette**
 Combining Machine Learning and Natural Language Processing for Automatic Keyword Extraction
 No 04-011 **Jayaweera, Prasad M.**
 A Unified Framework for e-Commerce Systems Development: *Business Process Patterns Perspective*
 No 04-013 **Söderström, Eva**
 B2B Standards Implementation: Issues and Solutions
 No 04-014 **Backlund, Per**

Development Process Knowledge Transfer through Method Adaptation, Implementation, and Use
 No 05-003 **Davies, Guy**
 Mapping and Integration of Schema Representations of Component Specifications
 No 05-004 **Jansson, Eva**
 Working Together when Being Apart – An Analysis of Distributed Collaborative Work through ICT from an Organizational and Psychosocial Perspective
 No 05-007 **Cöster, Rickard**
 Algorithms and Representations for Personalised Information Access
 No 05-009 **Ciobanu Morogan, Matei**
 Security System for Ad-hoc Wireless Networks based on Generic Secure Objects
 No 05-010 **Björck, Fredrik**
 Discovering Information Security Management
 No 05-012 **Brouwers, Lisa**
 Microsimulation Models for Disaster Policy Making
 No 05-014 **Näckros, Kjell**
 Visualising Security through Computer Games
 Investigating Game-Based Instruction in ICT Security: an Experimental approach
 No 05-015 **Bylund, Markus**
 A Design Rationale for Pervasive Computing
 No 05-016 **Strand, Mattias**
 External Data Incorporation into Data Warehouses
 No 05-020 **Casmir, Respickius**
 A Dynamic and Adaptive Information Security Awareness (DAISA) approach
 No 05-021 **Svensson, Harald**
 Developing Support for Agile and Plan-Driven Methods
 No 05-022 **Rudström, Åsa**
 Co-Construction of Hybrid Spaces
 No 06-005 **Lindgren, Tony**
 Methods of Solving Conflicts among Induced Rules
 No 06-009 **Wrigstad, Tobias**
 Owner-Based Alias Management
 No 06-011 **Skoglund, Mats**
 Curbing Dependencies in Software Evolution
 No 06-012 **Zdravkovic, Jelena**
 Process Integration for the Extended Enterprise
 No 06-013 **Olsson Neve, Theresia**
 Capturing and Analysing Emotions to Support Organisational Learning: The Affect Based Learning Matrix
 No 06-016 **Chaula, Job Asheri**
 A Socio-Technical Analysis of Information Systems Security Assurance
 A Case Study for Effective Assurance
 No 06-017 **Tarimo, Charles N.**
 ICT Security Readiness Checklist for Developing Countries:
 A Social-Technical Approach
 No 06-020 **Kifle Gelan, Mengistu**
 A Theoretical Model for Telemedicine
 - Social and Value Outcomes in Sub-Saharan Africa
 No 07-001 **Fernaesus, Ylva**

Let's Make a Digital Patchwork
 Designing for Children's Creative Play with Programming Materials
 No 07-003 **Bakari, Jabiri Kuwe**
 A Holistic Approach for Managing ICT Security in Non-Commercial Organisations
 A Case Study in a Developing Country
 No 07-004 **Sundholm, Hillevi**
 Spaces within Spaces: The Construction of a Collaborative Reality
 No 07-005 **Hansson, Karin**
 A Framework for Evaluation of Flood Management Strategies
 No 07-007 **Aidemark, Jan**
 Strategic Planning of Knowledge Management Systems
 - A Problem Exploration Approach
 No 07-009 **Jonsson, Martin**
 Sensing and Making Sense
 Designing Middleware for Context Aware Computing
 No 07-013 **Kabilan, Vandana**
 Ontology for Information Systems (O4IS) Design Methodology:
 Conceptualizing, Designing and Representing Domain Ontologies
 No 07-014 **Mattsson, Johan**
 Pointing, Placing, Touching
 - Physical Manipulation and Coordination Techniques for Interactive Meeting
 Spaces
 No 07-015 **Kessler, Anna-Maria**
 A Systemic Approach Framework for Operational Risk
 - SAFOR
 No 08-001 **Laaksolahti, Jarmo**
 Plot, Spectacle and Experience: Contributions to the design and evaluation of
 Interactive Storytelling
 No 08-002 **Van Nguyen Hong**
 Mobile Agent Approach to Congestion Control in Heterogeneous Networks
 No 08-003 **Rose-Mharie Åhlfeldt**
 Information Security in Distributed Healthcare
 - Exploring the Needs for Achieving Patient Safety and Patient Privacy
 No 08-004 **Sara Ljungblad**
 Beyond users:
 Grounding technology in experience
 No 08-005 **Eva Sjöqvist**
 Electronic Mail and its Possible Negative Aspects in Organizational Contexts
 No 08-006 **Thomas Sandholm**
 Statistical Methods for Computational Markets
 - Proportional Share Market Prediction and Admission Control
 No 08-007 **Lena Aggestam**
 IT-supported Knowledge Repositories:
 Increasing their Usefulness by Supporting Knowledge Capture
 No 08-008 **Jaana Nyfjord**
 Towards Integrating Agile Development and Risk Management
 No 08-009 **Åsa Smedberg**
 Online Communities and Learning for Health

- The Use of Online Health Communities and Online Expertise for People with Established Bad Habits
No 08-010 **Martin Henkel**
Service-based Processes
- Design for Business and Technology
No 08-012 **Jan Odelstad**
Many-Sorted Implicative Conceptual Systems
- No 09-001 **Marcus Nohlberg**
Securing Information Assets
- Understanding, Measuring and Protecting against Social Engineering Attacks
No 09-002 **Maria Håkansson**
Playing with Context
- Explicit and Implicit Interaction in Mobile Media Applications
No 09-003 **Petter Karlström**
Call of the Wild
Using language technology in the second language classroom
- No 09-009 **Ananda Edirisurya**
Design Support for e-Commerce Information Systems using Goal, Business and Process Modelling
- No 10-005 **Moses Niwe**
Organizational Patters for Knowledge Capture in B2B Engagements
- No 10-007 **Mats Wiklund**
Perception of Computer Games in Non-Gaming Contexts
- No 10-008 **Petra Sundström**
Designing Affective Loop Experiences
- No 10-009 **Tharaka Ilayperuma**
Improving E-Business Design through Business
- No 11-002 **David Sundgren**
The Apparent Arbitrariness of Second-Order Probability Distributions
- No 11-003 **Atelach Argaw**
Resource Lenient Approaches to Cross Language Information Retrieval using Amharic
- No 11-004 **Erik Perjons**
Model-Driven Networks, Enterprise Goals, Services and IT Systems
- No 11-005 **Lourino Chemane**
ICT Platform Integration – A MCDM Based Framework for the Establishment of Value Network
Case Study: Mozambique Government Electronic Network (GovNet)
- No 11-010 **Christofer Waldenström**
Supporting Dynamic Decision Making in Naval Search and Evasion Tasks
- No 11-012 **Gustaf Juell-Skielse**
Improving Organizational Effectiveness through Standard Application Packages and IT Services
- No 12-001 **Edephonc Ngemera Nfuka**
IT Governance in Tanzanian public sector organizations
- No 12-002 **Sumithra Velupillai**
Shades of Certainty
Annotation and Classification of Swedish Medical Records
- No 12-003 **Arvid Engström**

Collaborative Video Production After Television
 No12-007 **Fatima Jonsson**
 Hanging out in the game café.
 Contextualizing co-located game play practices and experiences
 No 12-008 **Mona Riabacke**
 A Prescriptive Approach to Eliciting Decision Information
 No 12-010 **Geoffrey Rwezaura Karokola**
 A Framework for Securing e-Government Services
 The Case of Tanzania
 No 13-001 **Evelyn Kigozi Kahiigi**
 A Collaborative E-learning Approach
 Exploring a Peer Assignment Review Process at the University Level in Uganda
 No 13-002 **Mattias Rost**
 Mobility is the message: Explorations in mobile media sharing
 No 13-003 **Rasika Dayarathna**
 Discovering Constructs and Dimensions for Information Privacy Metrics.
 No 13-004 **Magnus Johansson**
 Do Non Player Characters dream of electric sheep?
 No 13-006 **Baki Cakici**
 The Informed Gaze: On the Implications of ICT-Based Surveillance
 No 13-008 **Johan Eliasson**
 Tools for Designing Mobile Interaction with the Physical Environment in Outdoor Lessons
 No 13-010 **Andreas Nilsson**
 Doing IT Project Alignment – Adapting the DELTA Model using Design Science
 No. 14-001 **Ola Caster**
 Quantitative methods to support drug benefit-risk assessment
 No. 14-002 **Catarina Dudas**
 Learning from Multi-Objective Optimization of Production Systems - A method for analyzing solution sets from multi-objective optimization
 No. 14-003 **Thashmee M. Karunaratne**
 Learning predictive models from graph data using pattern mining
 No. 14-007 **David Hallberg**
 Lifelong learning - The social impact of digital villages as community resource centres on disadvantaged women
 No. 14-008 **Constantinos Giannoulis**
 Model-driven Alignment - Linking Business Strategy with Information Systems
 No. 14-014 **Jalal Nouri**
 Orchestrating Scaffolded Outdoor Mobile Learning Activities
 No. 14-016 **Jamie Walters**
 Distributed Immersive Participation – Realising Multi-Criteria Context-Centric Relationships on an Internet of Things
 No. 14-017 **Peter Mozelius**
 Education for All in Sri Lanka - ICT4D Hubs for Region-wide Dissemination of Blended Learning
 No. 15-001 **Maria Skeppstedt**
 Extracting Clinical Findings from Swedish Health Record Text
 No. 15-002 **Florence Kivunike**
 A Structured Approach for Evaluating the ICT Contribution to Development

No. 15-004 **Karin Hansson**

Accommodating differences – Power, belonging, and representation online

No. 15-008 **Thushani Weerasinghe**

Designing Online Courses for Individual and Collaborative Learning – A study of a virtual learning environment based in Sri Lanka

No. 15-009 **Tuwe Löfström**

On Effectively Creating Ensembles of Classifiers - Studies on Creation Strategies, Diversity and Predicting with Confidence

No. 15-010 **Shahzad Saleem**

Protecting the Integrity of Digital Evidence and Basic Human Rights During the Process of Digital Forensics