ENASE 2012

Proceedings of the
7th International Conference on
Evaluation of Novel Approaches to Software
Engineering

Wrocław, Poland

29 - 30 June, 2012

Edited by Joaquim Filipe and Leszek Maciaszek

BRIEF CONTENTS

INVITED SPEAKERS

Schahram Dustdar

Vienna University of Technology

Austria

Dimitris Karagiannis

University of Vienna

Austria

Steffen Staab

University of Koblenz-Landau

Germany

Pericles Loucopoulos

Loughborough University

U.K.

Yannis Manolopoulos

Aristotle University

Greece

SPECIAL SESSION CHAIR

SPECIAL SESSION ON MODEL-DRIVEN ARCHITECTURE AND MODELING-DRIVEN SOFTWARE DEVELOPMENT

Janis Osis, Riga Technical University, Latvia

ORGANIZING AND STEERING COMMITTEES

CONFERENCE CHAIR

Joaquim Filipe, Polytechnic Institute of Setúbal / INSTICC, Portugal

PROGRAM CHAIR

Leszek Maciaszek, Wroclaw University of Economics, Poland / Macquarie University ∼ Sydney, Australia

ORGANIZATION CHAIR

Helder Coelhas, INSTICC, Portugal

Artur Rot, Wroclaw University of Economics, Poland

PROCEEDINGS PRODUCTION

Helder Coelhas, INSTICC, Portugal

Andreia Costa, INSTICC, Portugal

Patrícia Duarte, INSTICC, Portugal

Bruno Encarnação, INSTICC, Portugal

Liliana Medina, INSTICC, Portugal

Raquel Pedrosa, INSTICC, Portugal

Vitor Pedrosa, INSTICC, Portugal

Cláudia Pinto, INSTICC, Portugal

CD-ROM PRODUCTION

Pedro Varela, INSTICC, Portugal

GRAPHICS PRODUCTION AND WEBDESIGNER

Daniel Pereira, INSTICC, Portugal

SECRETARIAT

Helder Coelhas, INSTICC, Portugal

WEBMASTER

Susana Ribeiro, INSTICC, Portugal

PROGRAM COMMITTEE

Guglielmo De Angelis, ISTI-CNR, Italy

Colin Atkinson, University of Mannheim, Germany

Costin Badica, University of Craiova, Romania

Ghassan Beydoun, University of Wollongong, Australia

Maria Bielikova, Slovak University of Technology in Bratislava, Slovak Republic

Piotr Bubacz, University of Zielona Gora, Poland

Dumitru Burdescu, University of Craiova, Romania

Wojciech Cellary, Poznan University of Economics, Poland

Rebeca Cortazar, University of Deusto, Spain

Massimo Cossentino, National Research Council, Italy

Bernard Coulette, Université De Toulouse Ii Le Mirail, France

Marcelo d'Amorim, Federal University of Pernambuco, Brazil

Philippe Dugerdil, Haute École de Gestion, University of Applied Sciences, Switzerland

Angelina Espinoza, Universidad Politécnica de Madrid, Spain

Joerg Evermann, Memorial University of Newfoundland, Canada

Maria João Ferreira, Universidade Portucalense, Portugal

Agata Filipowska, Poznan University of Economics, Poznan, Poland

Maria Ganzha, SRI PAS and University of Gdansk, Poland

Juan Garbajosa, Technical University of Madrid, UPM, Spain

Alessandro Garcia, Pontifical Catholic University of Rio de Janeiro, Brazil

Cesar Gonzalez-Perez, Institute of Heritage Sciences (Incipit), Spanish National Research Council (CSIC), Spain

Jeff Gray, University of Alabama, U.S.A.

Hans-Gerhard Gross, Delft University of Technology, The Netherlands

Rene Hexel, Griffith University, Australia

Emily Hill, Montclair State University, U.S.A.

Benjamin Hirsch, EBTIC / Khalifa University, U.A.E.

Charlotte Hug, Université Paris 1Panthéon-Sorbonne, France

Bernhard G. Humm, Hochschule Darmstadt - University of Applied Sciences, Germany

Zbigniew Huzar, Wroclaw University of Technology, Poland

Akira Imada, Brest State Technical University, Belarus

Stefan Jablonski, University of Bayreuth, Germany

Slinger Jansen, Utrecht University, The Netherlands

Monika Kaczmarek, Poznan University of Economics, Poland

Robert S. Laramee, Swansea University, U.K.

Xabier Larrucea, European Software Institute - Tecnalia, Spain

George Lepouras, University of Peloponnese, Greece

Francesca Lonetti, National Research Council (CNR) Pisa, Italy

Pericles Loucopoulos, Loughborough University, U.K.

Graham Low, University of New South Wales, Australia

Jian Lu, Nanjing University, China

André Ludwig, University of Leipzig, Germany

Ivan Lukovic, University of Novi Sad, Faculty of Technical Sciences, Serbia

Leszek Maciaszek, Wroclaw University of Economics, Poland / Macquarie University ∼ Sydney, Australia

Lech Madeyski, Wroclaw University of Technology, Poland

Leonardo Mariani, University of Milano Bicocca, Italy

Sascha Mueller-Feuerstein, Ansbach University of Applied Sciences, Germany

Johannes Müller, University of Leipzig, Germany

Anne Hee Hiong Ngu, Texas State University-San Marcos, U.S.A.

Andrzej Niesler, Wroclaw University of Economics, Poland

Janis Osis, Riga Technical University, Latvia

Marcin Paprzycki, Polish Academy of Science, Poland

Oscar Pastor, Universidad Politécnica de Valencia, Spain

Dana Petcu, West University of Timisoara, Romania

Naveen Prakash, MRCE, India

Lutz Prechelt, Freie Universität Berlin, Germany

Elke Pulvermueller, University of Osnabrueck, Germany

Rick Rabiser, Johannes Kepler University, Linz, Austria

Lukasz Radlinski, University of Szczecin, Poland

Gil Regev, Ecole Polytechnique Fédérale de Lausanne, Switzerland

Artur Rot, Wroclaw University of Economics, Poland

Radoslaw Rudek, Wroclaw University of Economics, Poland

Francisco Ruiz, Universidad de Castilla-La Mancha, Spain

Krzysztof Sacha, Warsaw University of Technology, Poland

Motoshi Saeki, Tokyo Institute of Technology, Japan

Sreedevi Sampath, UMBC, U.S.A.

Heiko Schuldt, University of Basel, Switzerland

Onn Shehory, IBM - Haifa Research Lab, Israel

Jerzy Surma, Warsaw School of Economics, Poland

Jakub Swacha, University of Szczecin, Poland

Stephanie Teufel, University of Fribourg, Switzerland

Rainer Unland, University of Duisburg-Essen, Germany

Olegas Vasilecas, Vilnius Gediminas Technical University, Lithuania

Krzysztof Wecel, Poznan University of Economics, Poland

Michael Whalen, University of Minnesota, U.S.A.

Igor Wojnicki, Agh University of Science and Technology, Poland

Kang Zhang, The University of Texas at Dallas, U.S.A.

AUXILIARY REVIEWERS

Fabiano Ferrari, Federal University of São Carlos, Brazil

Stefan Hanenberg, University of Duisburg-essen, Germany

Reinhold Kloos, University Duisburg-Essen, Germany

Ioan Lazar, Babes-Bolyai University, Romania

Otavio Lemos, UNIFESP, Brazil

SPECIAL SESSION PROGRAM COMMITTEE

SPECIAL SESSION ON MODEL-DRIVEN ARCHITECTURE AND MODELING-DRIVEN SOFTWARE DEVELOPMENT

Gundars Alksnis, Riga Technical University, Latvia

Erika Asnina, Riga Technical University, Latvia

Ruth Breu, Universität Innsbruck, Institut Für Informatik, Austria

Liliana Dobrica, University Politehnica of Bucharest, Romania

Jeff Gray, University of Alabama, U.S.A.

Petr Hnetynka, Charles University, Czech Republic

Tharaka Ilayperuma, University of Ruhuna, Sri Lanka

Lajos Kollár, University of Debrecen, Hungary

Christian Kop, Alpen-Adria-Universitaet Klagenfurt, Austria

Jose-Norberto Mazon, University of Alicante, Spain

Janis Osis, Riga Technical University, Latvia

Claus Pahl, Dublin City University, Ireland

Bazil Parv, Babes-Bolyai University, Romania

Raman Ramsin, Sharif University of Technology, Iran, Islamic Republic of

Janis Silins, Mapon SIA, Latvia

Jelena Zdravkovic, Stockholm University, Sweden

SELECTED PAPERS BOOK

A number of selected papers presented at ENASE 2012 will be published by Springer-Verlag in a CCIS Series book. This selection will be done by the Conference Chair and Program Chair, among the papers actually presented at the conference, based on a rigorous review by the ENASE 2012 Program Committee members.

FOREWORD

This volume contains the proceedings of the 7th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE) held in Wrocław, Poland, sponsored by the Institute for Systems and Technologies of Information, Control and Communication (INSTICC) and co-organized by the Wrocław University of Economics (WUE). The proceedings in this volume have been published by INSTICC in time for the staging of the conference. Moreover, starting from ENASE'2007, modified and extended versions of ENASE full papers are published as post-proceedings by Springer in Revised Selected Papers Series.

The previous six ENASE conferences took place in Erfurt, Germany (2006), Barcelona, Spain (2007), Madeira, Portugal (2008), Milan, Italy (2009), Athens, Greece (2010), and Beijing, China (2011). Three out of six past ENASE conferences were co-located with the International Conference on Enterprise Information Systems (ICEIS) and this trend is set to continue – the two conferences are going to be staged together again in 2013 in Angers, France.

There is a well-focused research community around the ENASE conferences. The mission of ENASE is to be a prime international forum to discuss and publish research findings and IT industry experiences with relation to evaluation of novel approaches to software engineering. By comparing novel approaches with established traditional practices and by evaluating them against software quality criteria, the ENASE conferences advance knowledge and research in software engineering, identify most hopeful trends and propose new directions for consideration by researchers and practitioners involved in large-scale software development and integration.

Overall, for the 7th ENASE in Wrocław we received 54 papers from 27 countries, of which 11 papers were accepted for publication and presentation as full papers. The papers were submitted for blind reviews to 82 renowned ENASE'2012 Program Committee members who registered in the ENASE conference system. The reviewing process was smoothly carried out by the great majority of the PC members. With multiple reviews for each paper, the final decision of acceptance/rejection was taken by the PC Chair Leszek Maciaszek. A few borderline papers were subjected to extra considerations and discussions before decisions were reached.

Statistically speaking, the ENASE'2012 acceptance ratio for full papers is 20.4% (based on 54 submissions), but considering that 9 papers were formally submitted in the category of "position" (rather than "regular") papers more truthful ratio for full papers is 24% (based on 45 submissions). This acceptance rate confirms the ENASE trend and the ENASE organizers' desire to ensure high quality of the conferences (all seven ENASE conferences had the acceptance rate above 30%).

At the conference, full papers (max. length of 10 pages) are presented in 30 minute sessions.

ENASE'2012 accepted also 12 short papers (max. 6 pages) for 20 minute presentations and 5 poster papers for display and discussions between the authors and the conference participants.

As a separate satellite event, ENASE'2012 hosts also the Special Session on Model-Driven Architecture and Modeling-Driven Software Development - MDA&MDSD'2012. Additionally the ENASE'2012 delegates have the benefit of participating in joint plenary sessions with ICEIS'2012 that includes keynote and panel presentations. The prominent list of keynotes consists of professors:

- Schahram Dustdar, Vienna University of Technology, Austria
- Dimitris Karagiannis, University of Vienna, Austria
- Steffen Staab, University of Koblenz-Landau, Germany
- Pericles Loucopoulos, Loughborough University, United Kingdom
- Yannis Manolopoulos, Aristotle University, Greece

We thank the authors and PC members for delivering, on both ends of the submissions/reviews spectrum, a great scientific content to the conference and to the proceedings. We thank the invited keynote and panel speakers. We thank the organizers and speakers of the special MDA&MDSD'2012 event.

The organization of this conference required a dedicated and coordinated effort of many people over an extended period of time. On the INSTICC side the main coordination was provided by Helder Coelhas and locally in Wrocław by Artur Rot, who was supported by a considerable number of staff members and students of the Institute of Business Informatics at WUE, in particular Wiesława Gryncewicz and Radosław Rudek.

We are confident that ENASE'2012 will be a great experience and we wish all delegates an exciting memorable time spent in Wrocław - the city that has gained the titles of 'the Meeting Place of Poland' and 'Venice of Poland', the co-host of Euro 2012 - European Football Championships, the 2016 European Capital of Culture (jointly with San Sebastian in Spain), with over a thousand years of history, the city where as many as 10 Nobel Prize winners lived, and which provides education to as many as 120 thousand students enrolled in the city's universities.

Joaquim Filipe
Polytechnic Institute of Setúbal / INSTICC, Portugal

Leszek Maciaszek
Wroclaw University of Economics, Poland / Macquarie University ∼ Sydney, Australia

CONTENTS

SPECIAL SESSION ON MODEL-DRIVEN ARCHITECTURE AND MODELING-DRIVEN SOFTWARE DEVELOPMENT

FULL **P**APERS

Refactoring Business Process Models
A Systematic Review

María Fernández-Ropero, Ricardo Pérez-Castillo and Mario Piattini

Instituto de Tecnologías y Sistemas de Información (ITSI), University of Castilla-La Mancha,
Paseo de la Universidad 4, 13071, Ciudad Real, Spain
{MariaS.Fernandez, Ricardo.PdelCastillo, Mario.Piattini}@uclm.es

Keywords: Refactoring, Business Process, Systematic Literature Review.

Abstract: Business processes are nowadays recognized as one of the intangible business assets that provide more competitive advantage to organizations. Organizations must therefore be able to manage their business process models and deal with their quality problems, i.e. lack of understandability, maintainability or reusability among others. Such quality problems are exacerbated in business processes models that were mined by reverse engineering from enterprise information systems, since business process are more likely to undergo inconsistencies, redundancies, etc. Refactoring has proved to be a suitable solution to cope with these quality problems. Refactoring changes the internal structure of a business process model while preserves its external behaviour. This paper presents an in-depth systematic review for collecting, categorizing and analyzing all the refactoring methods and techniques applied to business process models. The systematic review is conducted following the formal methodology proposed by *Kitchenhan*. The review reports 206 related studies, from which 16 were considered as primary studies. The most valuable conclusion is that none of these studies proposes refactoring techniques for business process models previously obtained by reverse engineering, which is considered as a greenfield research area.

1 INTRODUCTION

Business processes depict sequences of coordinated business activities as well as the involved roles and resources that organizations must carry out to achieve common goals (Weske, 2007). Business processes are today recognized as an intangible business asset that provides competitive advantages for organizations (Jeston et al., 2008). In addition, most business processes are automated by enterprise information systems.

To take an effective advantage of business process management, business processes need to be represented in models following standard notations such as BPMN (Business Process Modeling and Notation) (OMG, 2006). Organizations sometimes do not explicitly have their business processes models because they have never modeled their business processes before. Even when the organization has business process models, such models can be outdated or misaligned regarding the actual processes supported by enterprise information systems. In these cases, reverse engineering techniques can be used to obtain business process models from existing information systems (Pérez-Castillo et al., 2011a, Pérez-Castillo et al., 2011b). This way, the retrieved models are prone to have a lower quality degree since every reverse engineering technique is characterized by a semantic loss when the abstraction level is progressively increased (in this case from existing information systems to business processes).

In fact, bussiness process models can sometimes present quality faults such as redundancies, ambiguities, inconsistencies, lack of completeness, as well as non-adherence to conventions or standards, among others (Mens et al., 2007). Business process models particularly obtained by reverse engineering dramatically face these problems. For this reason its is necessary to carry out a refactoring process, which can solve the mentioned quality problems. Refactoring modifies the internal structure of business process models without changing or altering the external semantics. Refactoring techniques therefore make it possible to improve the quality of business processes, so that they become more understandable, maintenible and reusable (Dijkman et al., 2011).

The concept of refactoring was proposed by Opdyke in 1992 as a methodology for restructuring programs (Opdyke, 1992). In the last decades refactoring has emerged as a technique to improve the maintainability of software systems. Traditional refactoring primarily focuses on the level of source code for the software life cycle, but it can also be applied to the scope of the models (Mens and Tourwé, 2004). For example, refactoring has been applied to UML (Unified Modeling Language) models and also to business process modeling.

Refactoring is also frequently termed as restructuring since it is applied as the second stage of software modernization process. Software modernization advocates carrying out traditional software reengineering following the model-driven developments principles (i.e., considering all the involved artifacts as models at different abstraction levels). Software modernization consists of three stages: (i) reverse engineering, which represents the system at a higher level of abstraction; (ii) restructuring or refactoring, which represents the system at the same level of abstraction, improving one or more properties of the system, preserving its external behavior; and (iii) forward engineering, which generates the implementation of the system at a lower level of abstraction integrating the new features included in the previous stage.

Business process models refactoring, within a software modernization process, first needs to detect refactoring opportunities and then it applies refactoring operators. Currently, there are several techniques and approaches to refactor business process models. However, there is not a systematic review of all available techniques in the literature in order to understand its evolution and motivation in each case. For this reason this paper presents a systematic literature review to know the refactoring techniques and detect those techniques that can be particularly applied to business process models retrieved by reverse engineering. The review provides a summary of the state-of-the-art and identifies possible areas of research that have not been addressed yet.

The systematic literature review is planed and conducted following the process proposed by (Kitchenham and Charters, 2007) which is organized in: (1) planning, which defines the research questions and establishes the search protocol; (2) execution, in which the search of research studies is systematically carried out according to the protocol defined in the previous phase; and (3) analysis of results, in which some of the retrieved studies are considered as primary and are analyzed to draw conclusions.

After analyzing the studies retrieved during the search, the following main conclusions were obtained: (1) studies about business process refactoring hardly ever provide an empirical validation of their proposals, which reveals that this research field is not mature enough; (2) there is an increasing interest in the area in recent years; and (3) none of the studies proposes to refactor business process models previously obtained through reverse engineering from existing information systems. As a result, there is a potential research field that has not been addressed by now.

The remaining of the paper is organized as follows: Section 2 describes the first phase of the systematic review through the formulation of research questions. Section 3 corresponds to the second phase of the systematic review, the conduction. Section 4 describes the analysis of the results, concluding the process of *Kitchenham*. Finally, Section 5 shows the obtained conclusions after the execution of a systematic literature review.

2 PLANNING THE REVIEW

This section presents the planning of the review. It shows the research questions formulated and the development of the protocol to guide the review.

2.1 Research Questions

This section provides the research questions formulated in the review, which must be answered after analysing the obtained results. Table 1 shows the research questions RQ1 and the additional question AQ1.

2.2 Development of the Protocol

This section specifies the method used to carry out the systematic review. The review protocol provides guidelines to find primary studies that give an answer to the research questions. Thus, it is necessary to have some selection criteria of these primary studies for their inclusion in or exclusion from the systematic review. Then, a categorization of each primary study is performed through data extraction.

2.2.1 Formulation of the Search Strings

In order to formulate the search string it is necessary to know what to search and where to search.

The answer to the first question is obtained from search terms that are derived from the research questions ('Refactoring' and 'Business Process Model'), together with the related terms that are included in the search string by using the logical OR. Some of these terms are subsets of other terms. For this reason, it has been decided to select the most general terms in order to avoid redundancies. The search string is shown in Table 2.

As regards to where to search it is necessary to establish a series of digital libraries where the searches will be performed. The search of the string mentioned above is performed in the following digital libraries: (DL1) *ACM Digital Library*, (DL2) *Springer Link*, (DL3) *IEEE Xplore Digital Library* and (DL4) *Scopus*. Each of these digital libraries provides a different search, so it is necessary to make small changes in the search string in order to adapt it to each of these mechanisms.

2.2.2 Selection Criteria

This section defines the inclusion and exclusion criteria that must pass each of the studies retrieved by the search string in order to be considered as primary studies.

Inclusion criteria are those that determine if a study is considered for review or not (see Table 3), while exclusion criteria are applied after them to exclude non-relevant studies (see Table 4).

2.2.3 Study Selection Procedure

The procedure to retrieve primary studies for the systematic review is shown in Table 5.

3 CONDUCTING THE REVIEW

After planning the systematic literature review by the search protocol the primary studies are obtained by its execution. The steps for conducting the review are:

1. Adapt the search string for each digital library, since each one has different search engines.
2. Carry out a search in each of the digital libraries.
3. Apply de procedure of study selection to obtain the primary studies.
4. Use the data extraction mechanism for managing the relevant information from the studies.

Table 1: Research Questions.

Id.	Research Questions
RQ1	What techniques or procedures exist to carry out refactoring business process models?
AQ1	Are there any techniques or procedures to refactor business process models obtained through reverse engineering?

Table 2: Research String.

Id.	Research String
RS1	(Refactoring **OR** Restructuring **OR** Refactored **OR** Refactor) **AND** (Process Model **OR** BPMN **OR** Workflows **OR** Business Process)

Table 3: Inclusion Criteria.

Id.	Inclusion Criteria
IC1	The study answers directly to the research question
IC2	The study is focused on the refactoring of business process models
IC3	The study has been published between January 2006 and December 2011
IC4	The study provides empirical validation

Table 4: Exclusion Criteria.

Id.	Exclusion Criteria
EC1	The study has a business focus
EC2	The study has an approach away from software engineering
EC3	The study is duplicated
EC4	The study is written in a language different to English
EC5	The study shows only personal opinions or anecdotes without scientific basis

Table 5: Procedure for studies selection.

1. Read *Title* and *abstract* of study e_i
2. Apply Inclusion Criteria (IC)
   2.1. If $\forall\ c \in IC$, e_i SATISFIES c go to step 3
   2.2. If $\exists\ c \in IC$, e_i ¬SATISFIES c go to step 8
   2.3. If not enough information to determine it, go to step 3
3. Apply Exclusion Criteria (EC)
   3.1. If $\forall\ c \in EC$, e_i ¬SATISFIES c go to step 4
   3.2. If $\exists\ c \in EC$, e_i SATISFIES c go to step 8
   3.3. If not enough information to determine it, go to step 4
4. Read full text
5. Apply Inclusion Criteria again (IC)
   5.1. If $\forall\ c \in CI$, e_i SATISFIES c go to step 6
   5.2. If $\exists\ c \in CI$, e_i ¬SATISFIES c go to step 8
6. Apply Exclusion Criteria again (EC)
   6.1. If $\exists\ c \in CE$, e_i SATISFIES c go to step 8
   6.2. If $\forall\ c \in CE$, e_i ¬SATISFIES c go to step 7
7. Accept study
8. Refuse study

When applying the search protocol it is necessary to extract the relevant information from each retrieved study. To manage the information from each of the studies it has been prepared a *form of relevant information* that collects the relevant information for each retrieved study as an identifier of the study, the digital library where the study was recovered, the title of the study, the list of authors of the study and the year of publication of the study. Also, the form includes information relating to the satisfaction of the inclusion and exclusion criteria.

For each primary study, additional relevant information is also stored for later analysis, i.e. publication type (journal, conference or book), used notation (BPMN, Petri Nets, PTS Tree or others) and kind of empirical validation performed in the study (experiment, study case, poll, example or no validation). Moreover, the AQ1 (see Table 1) is formulated to each of the primary studies. This additional relevant information is collected to assess the quality of each study.

All of the primary studies are indexed in the database that provides the tool *EndNote* (Reuters, 2011). Additionally, all collected information during the execution of the systematic review is available online in (Pérez-Castillo and Fernández-Ropero, 2012).

4 ANALYSIS OF RESULTS

The obtained results are analyzed to be interpreted and to draw conclusions. The systematic review was carried out with two different analyses (see the following sections): (1) a descriptive statistical analysis and (2) a state-of-the-art assessment.

4.1 Statistics Analysis

After the execution of the search with the search string 206 studies were retrieved and after full execution of the review 16 studies were considered as primary studies. Table 6 provides some descriptive statistics.

Regarding the inclusion/exclusion criteria, only 7.77% of the studies met the inclusion and exclusion criteria proposed at the beginning. In the case of the *ACM* library only 28.57% of the studies were accepted as primary studies, with 4 studies published in journals, 9 published in conferences and 1 published in a book. In the case of *Springer* library only 4.17% of studies were accepted, with 1 published in a book. This decrease is due to the fact that most studies are duplicated in next libraries because those studies met the exclusion criterion

EC3. The same happened with *IEEE Xplore* library; for this reason none of studies were considered as primary studies from this library. Finally, from *Scopus* library only 0.96% of the retrieved studies were considered as primary studies, with 1 published in a journal, because it was the last library to search.

Concerning the kind of the publications, Table 6 shows that the majority of the studies (56%) have been published in conferences while 31% and 13% of the studies were respectively published in journals or books. This indicates that the subject under study does not have a great maturity degree.

Considering the year of publication, Figure 1 shows that there is a large increase of publications in the last years. In fact, in 2010 there was an increase in publication, but most of them are not relevant for this review since they did not meet the inclusion criteria (as answering the research questions) or since those studies did not provide empirical validation. Regarding primary studies, a growth is observed along the years with various studies per year from 2006.

Regarding the notation used to represent business process in the primary studies, Petri Nets predominates with 50% over all the primary studies (see Figure 2 (a)). The next most predominant notation is BPMN, which is used in 19% of the primary studies. Petri Nets might be more commonly used than BPMN given that Petri Nets is a well-proven notation and it has been used since the 60's while the BPMN notation is relatively new and it is not widespread.

With respect to the type of the empirical validation conducted in the primary studies, it shows that most studies (63%) provide only examples (see Figure 2 (b)). The reason for this could be a certain lack of maturity in the field and only proposals have been made.

Table 6: Summary of results.

Digital Library	Retrieved studies	Journal	Conference	Book	TOTAL	Primary/ Retrieved studies
		Primary studies				
ACM	49	4	9	1	14	28.57%
Springer	24	0	0	1	1	4.17%
IEEE	29	0	0	0	0	0.00%
Scopus	104	1	0	0	1	0.96%
TOTAL	206	5	9	2	16	7.77%

Additionally, the analysis of the primary studies revealed that no primary study considers business process models obtained through reverse engineering.
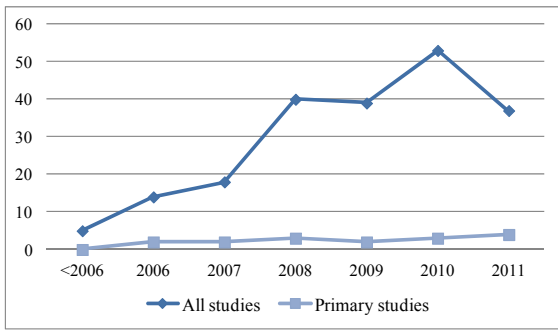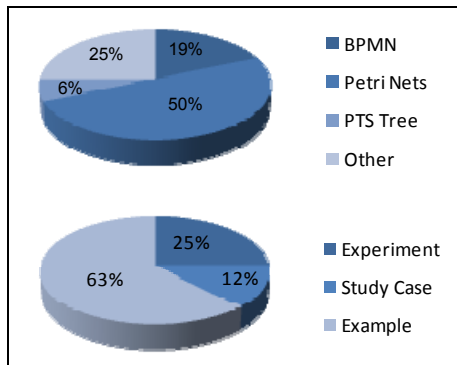
Figure 1: Distribution of year of publication.

Figure 2: Distribution of primary studies according (a) notation type and (b) empirical validation type.

Finally, the location of the authors involved in the studies was also analyzed, since it may be important

to understand the usefulness of the findings. At least ten countries were found, the most common of which being Germany (5 authors) as well as China and Switzerland (3 authors). Other countries as Austria, Netherlands, United Kingdom and United States have 2 authors, and Australia, Estonia and Japan have one author.

4.2 State-of-the-Art Analysis

After analyzing the whole set of primary studies, a set of common topics was collected. This analysis establishes a relationship between the most common, valuable topics and the primary studies addressing such topics.

Table 7 shows the reference of each primary study, that can be consulted in (Pérez-Castillo and Fernández-Ropero, 2012), and also shows the digital library where the study was retrieved and a list with all the topics indicating which topic appears or is addressed for each primary study.

There are 14 topics related to the studies. Topics are related to whether the study shows some scenarios (also called smells or refactoring opportunities) where refactoring would be necessary and, besides, if the algorithm used to make the detection is shown. Moreover, topics related to whether the study shows refactoring techniques and its algorithm. Other topic is whether the study proposes a metric. Furthermore, there are other topics related to quality as readability, reusability,

Table 7: Topics of primary studies.

Reference of primary study	Digital Library	Detect refactoring opportunities	Algorithms to detect opportunities	Measuring and Metrics	Refactoring techniques	Refactoring Algorithms	Business Process Readability	Business Process Understandability	Business Process Maintainability	Violations or Security	Business Process Reusability	Implementation (Tools)	Software Processes or Workflow	PAIS (Process-Aware Information System)	Process Variants models
(Dijkman et al., 2011)	DL1	◆	◆	◆				◆	◆		◆				◆
(Dumas et al., 2011)	DL1	◆	◆				◆		◆			◆			
(Weber et al., 2011)	DL1	◆			◆		◆		◆			◆		◆	◆
(Zeng et al., 2010)	DL1	◆	◆		◆	◆								◆	◆
(Hanakawa, 2011)	DL1				◆				◆				◆		
(Leopold et al., 2010)	DL1	◆	◆		◆			◆	◆						
(Feineman, 2010)	DL4	◆	◆					◆	◆			◆			
(Weber and Reichert, 2008)	DL1	◆			◆			◆	◆					◆	◆
(Awad et al., 2009)	DL1	◆			◆	◆				◆					
(Koehler et al., 2008)	DL1	◆			◆			◆	◆		◆	◆			
(Yan and Wang, 2009)	DL1	◆													
(Vanhatalo et al., 2008)	DL1	◆			◆	◆		◆							
(Chivers and McDermid, 2006)	DL1				◆					◆			◆		
(Wang et al., 2007)	DL1	◆			◆										
(Küster et al., 2006)	DL2	◆			◆										
(Singh et al., 2007)	DL1				◆								◆		

144

security, understandability and maintainability. A topic about whether the study proposes the implementation of its algorithms or not. There is a topic about whether the study is generic as software process or workflow but it can be applied in this context. Also, there is a topic on whether the study is about PAIS (Process-Aware Information System). Finally, there is a topic to indicate whether the study makes refactoring of process variants models.

5 CONCLUSIONS

The paper presents a systematic literature review of refactoring techniques and methods to be applied to business process models obtained by reverse engineering. The review has been carried out by following the formal methodology proposed by *Kitchenham*.

In total, 206 relevant studies were found in four different digital libraries (*ACM*, *SpringerLink*, *IEEE Xplore* and *Scopus*). 16 of these studies were considered as primary according to the inclusion and exclusion criteria and specific data were collected from them in order to analyze them and to obtain conclusions. After applying a statistical analysis the most valuable findings were the following: (1) as negative aspects: little empirical validation performed. Most of the studies considered only an example of the techniques or methods proposed, and some were proposed as future work to validate their proposals through study cases; (2) as beneficial aspects: growing interest in the field due to increased studies in recent years. It is also an area of research which is not mature enough, so it is interesting to address it.

Particularly, refactoring techniques have not been especially developed to business processes obtained by reverse engineering. Therefore, it may be a possible field in which to make further research efforts since it has not been addressed yet by the research community.

ACKNOWLEDGEMENTS

REFERENCES

Dijkman, R., Gfeller, B., Küster, J. & Völzer, H. 2011. Identifying refactoring opportunities in process model repositories. *Information and Software Technology*.

Jeston, J., Nelis, J. & Davenport, T. 2008. Business Process Management: Practical Guidelines to Successful Implementations, NV, USA, Butterworth-Heinemann (Elsevier Ltd.).

Kitchenham, B. & Charters, S. 2007. Guidelines for performing systematic literature reviews in software engineering. *Engineering*, 2.

Mens, T., Taentzer, G. & Müller, D. Year. Challenges in model refactoring. In, 2007.

Mens, T. & Tourwé, T. 2004. A survey of software refactoring. Software Engineering, *IEEE Transactions on*, 30, 126-139.

Omg. 2006. Business Process Modeling Notation Specification 1.0 [Online]. Available: http://www.omg.org/bpmn/Documents/OMG_Final_Adopted_BPMN_1-0_Spec_06-02-01.pdf [Accessed].

Opdyke, W. F. 1992. Refactoring: A program restructuring aid in designing object-oriented application frameworks. PhD thesis, University of Illinois at Urbana-Champaign.

Pérez-Castillo, R. & Fernández-Ropero, M. 2012. Refactoring Business Process Models - A Systematic Review [Online]. Available: http://alarcos.esi.uclm.es/per/rpdelcastillo/SLR.html [Accessed 21/02/2012 2012].

Pérez-Castillo, R., Fernández-Ropero, M., Guzmán, I. G.-R. D. & Piattini, M. 2011a. MARBLE. A Business Process Archeology Tool. *27th IEEE International Conference on Software Maintenance (ICSM 2011)*. Williamsburg, VI.

Pérez-Castillo, R., García-Rodríguez De Guzmán, I. & Piattini, M. 2011b. Business Process Archeology using MARBLE. *Information and Software Technology*.

Reuters, T. 2011. EndNote ®. Bibliographies Made Easy ™ http://www.endnote.com/.

Weske, M. 2007. Business Process Management: Concepts, Languages, Architectures, Leipzig, Alemania, *Springer-Verlag Berlin Heidelberg*.