

ICEIS 2012



14th International Conference on Enterprise Information Systems

Table of Contents

Introduction

Keynote Speakers

Volume 1 - Databases and Information Systems Integration / Artificial Intelligence and Decision Support Systems

Volume 2 - Information Systems Analysis and Specification / Software Agents and Internet Computing

Volume 3 - Human-Computer Interaction / Enterprise Architecture

Workshops

WOSIS, MSVVEIS/WEBI, NLPCS and IDEE

Special Sessions

SCOE, NTMIST and MDDIS

Wrocław, Poland
28 June - 1 July, 2012

<http://www.iceis.org>

 iceis.secretariat@insticc.org

Sponsored by:



In Cooperation with:



Co-organized by:



Proceedings will be submitted for indexing by:



ENASE 2012

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

Wrocław, Poland

29 - 30 June, 2012

<http://www.enase.org>


 enase.secretariat@insticc.org

Table of Contents

Introduction

Keynote Speakers

Volume 1

Special Session

MDA&MDSD - Model-Driven Architecture and Modeling-Driven Software Development

Sponsored by:



Co-organized by:



Proceedings will be submitted for indexing by:



ICEIS 2012

Proceedings of the
14th International Conference on
Enterprise Information Systems

Volume 1

Wrocław, Poland

28 June - 1 July, 2012

Sponsored by
**INSTICC – Institute for Systems and Technologies of Information, Control
and Communication**

Co-organized by
Wrocław University of Economics

In Cooperation with
AAAI – Association for the Advancement of Artificial Intelligence
**IEICE – The Institute of Electronics, Information and Communication
Engineers**
SWIM – IEICE Special Interest Group on Software Enterprise Modelling
ACM – Association for Computing Machinery
**ACM SIGMIS – ACM Special Interest Group on Management Information
Systems**
**ACM SIGCHI – ACM Special Interest Group on Computer Human
Interaction**

Copyright © 2012 SciTePress – Science and Technology Publications
All rights reserved

Edited by Leszek Maciaszek, Alfredo Cuzzocrea and José Cordeiro

Printed in Portugal

ISBN: 978-989-8565-10-5

Depósito Legal: 344492/12

<http://www.iceis.org/>
iceis.secretariat@insticc.org

BRIEF CONTENTS

INVITED SPEAKERS	IV
ORGANIZING AND STEERING COMMITTEES	V
SENIOR PROGRAM COMMITTEE	VI
PROGRAM COMMITTEE	VII
AUXILIARY REVIEWERS	XIII
SELECTED PAPERS BOOK	XIII
FOREWORD	XV
CONTENTS	XVII

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

ORGANIZING AND STEERING COMMITTEES

CONFERENCE CHAIR

Joaquim Filipe, Polytechnic Institute of Setúbal / INSTICC, Portugal

PROGRAM CO-CHAIRS

José Cordeiro, Polytechnic Institute of Setúbal / INSTICC, Portugal

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

Alfredo Cuzzocrea, ICAR-CNR and University of Calabria, Italy

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

José Varela, INSTICC, Portugal

CD-ROM PRODUCTION

Pedro Varela, INSTICC, Portugal

GRAPHICS PRODUCTION AND WEBDESIGNER

Daniel Pereira, INSTICC, Portugal

SECRETARIAT

Vitor Pedrosa, INSTICC, Portugal

WEBMASTER

Susana Ribeiro, INSTICC, Portugal

SENIOR PROGRAM COMMITTEE

Balbir Barn, Middlesex University, U.K.

Albert Cheng, University of Houston, U.S.A.

Jan Dietz, Delft University of Technology, The Netherlands

Schahram Dustdar, Vienna University of Technology, Austria

António Figueiredo, University of Coimbra, Portugal

Michel Léonard, CUI, University of Geneva, Switzerland

Kecheng Liu, University of Reading, U.K.

Pericles Loucopoulos, Loughborough University, U.K.

Andrea de Lucia, Università degli Studi di Salerno, Italy

Yannis Manolopoulos, Aristotle University, Greece

Masao Johannes Matsumoto, Solution Research Lab, Japan

Alain Pirotte, University of Louvain, Belgium

Klaus Pohl, University of Duisburg-Essen, Germany

Matthias Rauterberg, Eindhoven University of Technology, The Netherlands

Colette Rolland, Université Paris 1 Panthéon-Sorbonne, France

Narczyz Roztocki, State University of New York at New Paltz, U.S.A.

Abdel-Badeeh Mohamed Salem, Ain Shams University, Egypt

Bernadette Sharp, Staffordshire University, U.K.

Alexander Smirnov, SPIIRAS, Russian Academy of Sciences, Russian Federation

Ronald Stamper, Measur Ltd, U.K.

Merrill Warkentin, Mississippi State University, U.S.A.

PROGRAM COMMITTEE

Miguel Angel Martinez Aguilar, University of Murcia, Spain

Patrick Albers, ESEO - Ecole Supérieure D'Electronique de L'Ouest, France

Rainer Alt, University of Leipzig, Germany

Vasco Amaral, CITI FCT/UNL, Portugal

Andreas S. Andreou, Cyprus University of Technology, Cyprus

Wudhichai Assawinchaichote, King Mongkut's University of Technology Thonburi, Thailand

Cecilia Baranauskas, State University of Campinas - Unicamp, Brazil

Senén Barro, University of Santiago de Compostela, Spain

Rémi Bastide, ISIS - CUFR Jean-François Champollion, France

Bernhard Bauer, University of Augsburg, Germany

Lamia Hadrich Belguith, ANLP Research Group, MIRACL, University of Sfax, Tunisia

Jorge Bernardino, Institute Polytechnic of Coimbra - ISEC, Portugal

Felix Biscarri, University of Seville, Spain

Oliver Bittel, HTWG Konstanz - University of Applied Sciences, Germany

Danielle Boulanger, IAE- Université Jean Moulin Lyon 3, France

Jean-Louis Boulanger, CERTIFER, France

Peter Busch, Macquarie University ~ Sydney, Australia

Miguel Calejo, Declarativa, Portugal

Coral Calero, University of Castilla - La Mancha, Spain

Luis M. Camarinha-Matos, New University of Lisbon, Portugal

Olivier Camp, ESEO, France

Roy Campbell, University of Illinois at Urbana-Champaign, U.S.A.

Gerardo Canfora, University of Sannio, Italy

Manuel Isidoro Capel-Tuñón, University of Granada, Spain

Angélica Caro, University of Bio-Bio, Chile

Jose Jesus Castro-schez, Escuela Superior de Informatica, Spain

Luca Cernuzzi, Universidad Católica "Nuestra Señora de la Asunción", Paraguay

Sergio de Cesare, Brunel University, U.K.

Ming-Puu Chen, National Taiwan Normal University, Taiwan

Shiping Chen, CSIRO ICT Centre Australia, Australia

Shu-Ching Chen, Florida International University, U.S.A.

Max Chevalier, Institut de Recherche en Informatique de Toulouse UMR 5505, France

Witold Chmielarz, Warsaw University, Poland

Daniela Barreiro Claro, Universidade Federal da Bahia (UFBA), Brazil

Cesar Collazos, Universidad del Cauca, Colombia

Jose Eduardo Corcoles, Castilla-La Mancha University, Spain

Antonio Corral, University of Almeria, Spain

Karl Cox, University of Brighton, U.K.

Sharon Cox, Birmingham City University, U.K.

Alfredo Cuzzocrea, ICAR-CNR and University of Calabria, Italy

Maria Damiani, University of Milan, Italy

Vincenzo Deufemia, Università di Salerno, Italy

Kamil Dimililer, Near East University, Cyprus

José Javier Dolado, University of the Basque Country, Spain

Dulce Domingos, Faculty of Science - University of Lisbon, Portugal

César Domínguez, Universidad de La Rioja, Spain

Ming Dong, Shanghai Jiao Tong University, China

PROGRAM COMMITTEE (CONT.)

Juan C. Dueñas, Universidad Politécnica de Madrid, Spain

Hans-Dieter Ehrich, Technische Universitaet Braunschweig, Germany

João Faria, FEUP - Faculty of Engineering of the University of Porto, Portugal

Antonio Fariña, University of A Coruña, Spain

Jamel Feki, University of Sfax - Faculté Des Sciences Economiques et de Gestion de Sfax, Tunisia

Antonio Fernández-Caballero, Universidad de Castilla-la Mancha, Spain

Edilson Ferneda, Catholic University of Brasília, Brazil

Maria João Silva Costa Ferreira, Universidade Portucalense, Portugal

Paulo Ferreira, INESC-ID / IST, Portugal

Filomena Ferrucci, Università di Salerno, Italy

Rita Francese, Università degli Studi di Salerno, Italy

Bogdan Franczyk, University of Leipzig, Germany

Ana Fred, Technical University of Lisbon / IT, Portugal

Lixin Fu, University of North Carolina, Greensboro, U.S.A.

Mariagrazia Fugini, Politecnico di Milano, Italy

Jose A. Gallud, University of Castilla-la Mancha, Spain

Matjaz Gams, Jozef Stefan Institute, Slovenia

Maria Ganzha, SRI PAS and University of Gdansk, Poland

Juan Garbajosa, Technical University of Madrid, Spain

Mouzhi Ge, Universitaet der Bundeswehr Munich, Germany

Marcela Genero, University of Castilla-La Mancha, Spain

Joseph Giampapa, Carnegie Mellon University, U.S.A.

Raúl Giráldez, Pablo de Olavide University of Seville, Spain

Pascual Gonzalez, Universidad de Castilla-la Mancha, Spain

Robert Goodwin, Flinders University of South Australia, Australia

Feliz Gouveia, University Fernando Pessoa / Cerem, Portugal

Luis Borges Gouveia, Universidade Fernando Pessoa, Portugal

Virginie Govaere, INRS, France

Janis Grabis, Riga Technical University, Latvia

Maria Carmen Penadés Gramaje, Universitat Politècnica de València, Spain

Sven Groppe, University of Lübeck, Germany

Wieslawa Gryncewicz, Wroclaw University of Economics, Poland

Nuno Guimarães, Lasige / Faculty of Sciences, University of Lisbon, Portugal

Maki K. Habib, The American University in Cairo, Egypt

Yaakov Hacoheh-Kerner, Jerusalem College of Technology (Machon Lev), Israel

Sylvain Hallé, Université du Québec à Chicoutimi, Canada

Slimane Hammoudi, ESEO, France

Wahab Hamou-Lhadj, Concordia University, Canada

Christian Heinlein, Aalen University, Germany

Markus Helfert, Dublin City University, Ireland

Suvineetha Herath, Richard Stockton State College of New Jersey, U.S.A.

Orland Hoerber, Memorial University of Newfoundland, Canada

Wladyslaw Homenda, Warsaw University of Technology, Poland

Jun Hong, Queen's University Belfast, U.K.

Wei-Chiang Hong, Oriental Institute of Technology, Taiwan

PROGRAM COMMITTEE (CONT.)

Miguel J. Hornos, University of Granada, Spain

Kai-I Huang, Tunghai University, Taiwan

Akram Idani, Grenoble INP - Grenoble Institute of Technology, France

Marta Indulska, The University of Queensland, Australia

François Jacquenet, University of Saint-Étienne, France

Arturo Jaime, Universidad de La Rioja, Spain

Marijn Janssen, Delft University of Technology, The Netherlands

Wassim Jaziri, Higher Institute of Computer and Multimedia of Sfax, Tunisia

Paul Johannesson, Royal Institute of Technology, Sweden

Jan Jürjens, TU Dortmund & Fraunhofer ISST, Germany

Michail Kalogiannakis, University of Crete, Greece

Nikos Karacapilidis, University of Patras, Greece

Nikitas Karanikolas, Technological Educational Institute of Athens (TEI-A), Greece

Stamatis Karnouskos, SAP, Germany

Andrea Kienle, University of Applied Sciences, Dortmund, Germany

Marite Kirikova, Riga Technical University, Latvia

Alexander Knapp, Universität Augsburg, Germany

Fotis Kokkoras, TEI of Larisa, Greece

Ryszard Kowalczyk, Swinburne University of Technology, Australia

John Krogstie, NTNU, Norway

Subodha Kumar, Texas A&M University, U.S.A.

Rob Kusters, Eindhoven University of Technology & Open University of the Netherlands, The Netherlands

Halina Kwasnicka, Wroclaw University of Technology, Poland

Alain Leger, France Telecom Orange Labs, France

Kauko Leiviskä, University of Oulu, Finland

Daniel Lemire, UQAM - University of Quebec at Montreal, Canada

Da-Yin Liao, Applied Wireless Identifications, U.S.A.

Luis Jiménez Linares, University of de Castilla-La Mancha, Spain

Panos Linos, Butler University, U.S.A.

Stephane Loiseau, Leria, France

João Correia Lopes, Faculdade de Engenharia da Universidade do Porto/INESC Porto, Portugal

Maria Filomena Cerqueira de Castro Lopes, Universidade Portucalense Infante D. Henrique, Portugal

María Dolores Lozano, University of Castilla-la Mancha, Spain

Miguel R. Luaces, Universidade da Coruña, Spain

André Ludwig, University of Leipzig, Germany

Vicente Luque-Centeno, Carlos III University of Madrid, Spain

Mark Lycett, Brunel University, U.K.

Lukasz Lysik, Wroclaw University of Economics, Poland

Cristiano Maciel, Universidade Federal de Mato Grosso, Brazil

Rita Suzana Pitangueira Maciel, Federal University of Bahia, Brazil

Mirko Malekovic, University of Zagreb, Croatia

Nuno Mamede, INESC-ID, Portugal

Paolo Maresca, Università Federico II, Italy

Pierre Maret, Université de Saint Etienne, France

Tiziana Margaria, University of Potsdam, Germany

Farhi Marir, London Metropolitan University, U.K.

Herve Martin, Grenoble University, France

Maria João Martins, Instituto Superior Tecnico, Portugal

PROGRAM COMMITTEE (CONT.)

Katsuhisa Maruyama, Ritsumeikan University, Japan

Viviana Mascardi, University of Genoa, Computer Science Department, Italy

David Martins de Matos, L2F / INESC-ID Lisboa / Instituto Superior Técnico, Portugal

Wolfgang Mayer, University of South Australia, Australia

Javier Medina, University of Granada, Spain

Jerzy Michnik, University of Economics in Katowice, Poland

Luo Ming, Southeastern Institute of Manufacturing and Technology, Singapore

Michele Missikoff, IASI-CNR, Italy

Ghodrat Moghadampour, Vaasa University of Applied Sciences, Finland

Pascal Molli, LINA, University of Nantes, France

Lars Mönch, FernUniversität in Hagen, Germany

Valérie Monfort, SOIE Tunis, Tunisia

Francisco Montero, University of Castilla-la Mancha, Spain

Carlos León de Mora, University of Seville, Spain

Paula Morais, Portucalense University, Portugal

Fernando Moreira, Universidade Portucalense, Portugal

Haralambos Mouratidis, University of East London, U.K.

Pietro Murano, University of Salford, U.K.

Ana Neves, knowman - Consultadoria em Gestão, Lda, Portugal

Matthias Nickles, Technical University of Munich, Germany

Ann Nosseir, British University in Egypt, Egypt

Jose Angel Olivas, Universidad de Castilla - La Mancha, Spain

David Olson, University of Nebraska, U.S.A.

Mehmet Orgun, Macquarie University ~ Sydney, Australia

Andrés Muñoz Ortega, Catholic University of Murcia (UCAM), Spain

Samia Oussena, University of West London, U.K.

Sietse Overbeek, Delft University of Technology, The Netherlands

Tansel Ozyer, TOBB ETU, Turkey

Claus Pahl, Dublin City University, Ireland

Marcin Paprzycki, Polish Academy of Science, Poland

José R. Paramá, University of A Coruña, Spain

Eric Pardede, La Trobe University, Australia

Viviana Patti, University of Torino, Italy

Loris Penserini, FBK-IRST, Italy

Massimiliano Di Penta, University of Sannio, Italy

Laurent Péridy, IMA-UCO, France

Dana Petcu, West University of Timisoara, Romania

Yannis A. Phillis, Technical University of Crete, Greece

Josef Pieprzyk, Macquarie University, Australia

Selwyn Piramuthu, University of Florida, U.S.A.

José Pires, Escola Superior de Tecnologia e Gestão, Portugal

Luís Ferreira Pires, University of Twente, The Netherlands

Geert Poels, Ghent University, Belgium

Michal Polasik, Nicolaus Copernicus University, Poland

Srini Ramaswamy, ABB, India

T. Ramayah, Universiti Sains Malaysia, Malaysia

Pedro Ramos, Instituto Superior das Ciências do Trabalho e da Empresa, Portugal

Marek Reformat, University of Alberta, Canada

Francisco Regateiro, Instituto Superior Técnico, Portugal

Hajo A. Reijers, Eindhoven University of Technology, The Netherlands

PROGRAM COMMITTEE (CONT.)

Ulrich Reimer, University of Applied Sciences St. Gallen, Switzerland

Marinette Revenu, Greyc Ensicaen, France

Nuno de Magalhães Ribeiro, Universidade Fernando Pessoa, Portugal

Debbie Richards, Macquarie University ~ Sydney, Australia

Michele Risi, University of Salerno, Italy

David Rivreau, Université Catholique de L'ouest, France

Alfonso Rodriguez, University of Bio-Bio, Chile

Daniel Rodriguez, University of Alcalá, Spain

Pilar Rodriguez, Universidad Autónoma de Madrid, Spain

Oscar Mario Rodriguez-Elias, Institute of Technology of Hermosillo, Mexico

Erik Rolland, University of California at Riverside, U.S.A.

Jose Raul Romero, University of Cordoba, Spain

David G. Rosado, University of Castilla-la Mancha, Spain

Gustavo Rossi, Lifia, Argentina

Artur Rot, Wroclaw University of Economics, Poland

Francisco Ruiz, Universidad de Castilla-La Mancha, Spain

Roberto Ruiz, Pablo de Olavide University, Spain

Belen Vela Sanchez, Rey Juan Carlos University, Spain

Luis Enrique Sánchez, Sicaman Nuevas Tecnologias S.L., Spain

Manuel Filipe Santos, University of Minho, Portugal

Jurek Sasiadek, Carleton University, Canada

Andrea Schaerf, Università di Udine, Italy

Daniel Schang, ESEO, France

Sissel Guttormsen Schär, Institute for Medical Education, Switzerland

Isabel Seruca, Universidade Portucalense, Portugal

Jianhua Shao, Cardiff University, U.K.

Hossein Sharif, University of Portsmouth, U.K.

Alberto Silva, INESC, Portugal

Sean Siqueira, Federal University of the State of Rio de Janeiro, Brazil

Spiros Sirmakessis, Technological Educational Institution of Messolongi, Greece

Hala Skaf-molli, Nantes University, France

Chantal Soule-Dupuy, Université Toulouse 1, France

José Neuman de Souza, Universidade Federal do Ceará, Brazil

Martin Stanton, Manchester Metropolitan University, U.K.

Chris Stary, University of Linz, Austria

Dick Stenmark, Gothenburg University, Sweden

Stefan Strecker, University of Hagen, Germany

Vijayan Sugumaran, Oakland University, U.S.A.

Hiroki Suguri, Miyagi University, Japan

Lily Sun, University of Reading, U.K.

Raj Sunderraman, Georgia State University, U.S.A.

Jerzy Surma, Warsaw School of Economics, Poland

Gion K. Svedberg, Malmo University, Sweden

Ryszard Tadeusiewicz, AGH University of Science and Technology, Poland

Vladimir Tarasov, Jönköping University, Sweden

Arthur Tatnall, Victoria University, Australia

Sotirios Terzis, University of Strathclyde, U.K.

Claudine Toffolon, Université du Maine, France

Theodoros Tzouramanis, University of the Aegean, Greece

Athina Vakali, Aristotle University, Greece

José Ângelo Braga de Vasconcelos, Universidade Fernando Pessoa, Portugal

PROGRAM COMMITTEE (CONT.)

Michael Vassilakopoulos, University of Central Greece, Greece

Christine Verdier, LIG - University Joseph Fourier Grenoble, France

François Vernadat, European Court of Auditors, France

Maria Esther Vidal, Universidad Simon Bolivar, Venezuela

Aurora Vizcaino, Escuela Superior de Informática, Spain

Bing Wang, University of Hull, U.K.

Dariusz Wawrzyniak, Wroclaw University of Economics, Poland

Hans Weghorn, BW Cooperative State University Stuttgart, Germany

Hans Weigand, Tilburg University, The Netherlands

Gerhard Weiss, University of Maastricht, The Netherlands

Graham Winstanley, University of Brighton, U.K.

Wita Wojtkowski, Boise State University, U.S.A.

Viacheslav Wolfengagen, Institute JurInfoR, Russian Federation

Andreas Wombacher, University of Twente, The Netherlands

Robert Wrembel, Poznan University of Technology, Poland

Stanislaw Wrycza, University of Gdansk, Poland

Min Wu, Oracle, U.S.A.

Wen-Yen Wu, I-Shou University, Taiwan

Mudasser Wyne, National University, U.S.A.

Haiping Xu, University of Massachusetts Dartmouth, U.S.A.

Sadok Ben Yahia, Faculty of Sciences of Tunis, Tunisia

Lili Yang, Loughborough University, U.K.

Ping Yu, University of Wollongong, Australia

Yugang Yu, Erasmus University, The Netherlands

Wei Zhou, ESCP Europe, France

Eugenio Zimeo, University of Sannio, Italy

Lin Zongkai, Chinese Academy of Sciences, China

AUXILIARY REVIEWERS

Ankica Barisic, FCT UNL, Portugal

Bruno Barroca, CITI FCT UNL, Portugal

Gabriele Bavota, University of Salerno, Italy

Paulo Carreira, University of Lisbon, Portugal

Jan Claes, Ghent University, Belgium

Rui Domingues, FCT/UNL, Portugal

Jessica Diaz Fernandez, Universidad Politécnica de Madrid (Technical U. of Madrid), Spain

Daniel Lopez Fernandez, UPM, Spain

Rodrigo Garcia-Carmona, Universidad Politécnica de Madrid, Spain

Laura Sánchez González, University of Castilla La Mancha, Spain

Krzysztof Kania, University of Economics in Katowice, Poland

Sandra Lovrencic, University of Zagreb, Faculty of organization and informatics Varazdin, Croatia

Paloma Cáceres García de Marina, Rey Juan Carlos University, Spain

Luã Marcelo Muriana, UFMT, Brazil

David Musat, UPM, Italy

Álvaro Navas, Universidad Politécnica de Madrid, Spain

Annibale Panichella, University of Salerno, Italy

Ignazio Passero, University of Salerno, Italy

Jonas Poelmans, Faculty of Business and Economics, K.U. Leuven, Belgium

Hércules Antonio do Prado, Embrapa/Universidade Católica de Brasília, Brazil

Abdallah Qusef, University of Salerno, Italy

Federica Sarro, Università di Salerno, Italy

Markus Schatten, University of Zagreb, Croatia

Claudio Schifanella, Università degli Studi di Torino, Italy

Diego Seco, University of A Coruña, Spain

SELECTED PAPERS BOOK

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

Database Schema Elicitation to Modernize Relational Databases

Ricardo Pérez-Castillo¹, Ignacio García Rodríguez de Guzmán¹, Danilo Caivano² and Mario Piattini¹

¹*Instituto de Tecnologías y Sistemas de Información (ITSI), University of Castilla-La Mancha,
Paseo de la Universidad 4, 13071, Ciudad Real, Spain*

²*Department of Informatics, University of Bari, Via E. Orabona, 4, 70126 Bari, Italy
{ricardo.pdelcastillo, ignacio.grodriguez, mario.piattini}@uclm.es, caivano@di.uniba.it*

Keywords: Database Modernization, Legacy Systems, ADM, KDM, SQL, Metamodel, Model Transformations, QVT.

Abstract: Legacy enterprise systems mainly consist of two kinds of artefacts: source code and databases. Typically, the maintenance of those artefacts is carried out through re-engineering processes in isolated manners. However, for a more effective maintenance of the whole system both should be analysed and evolved jointly according to ADM (Architecture-Driven Modernization) approach. Thus, the ROI and the lifespan of the legacy system are expected to improve. In this sense, this paper proposes the schema elicitation technique for recovering the relational database schema that is minimally used by the source code. For this purpose, the technique analyses database queries embedded in the legacy source code in order to remove the dead parts of the database schema. Also, this proposal has been validated throughout a real-life case study.

1 INTRODUCTION

Today, many organizations have huge legacy systems supported by relational databases (Blaha, 2001), and these systems are not immune to software ageing (Visaggio, 2001). Nevertheless, the erosion not only affects the source code, but also databases age gradually. For instance, in order to adapt the system to new requirements, new tables and/or columns are added to the schema of the database; other tables are modified and even discarded without erasing them from the database; and so on. These changes over time generate problems related to inconsistency, redundancy and integrity among others.

Therefore, organizations must address maintenance processes in their legacy information systems. The entire replacement of these systems has a great impact in technological and economic terms (Sneed, 2005). So that, maintenance based on evolutionary reengineering processes has typically been carried out (Bianchi et al., 2003). Moreover, the typical re-engineering process has been shifted to the so-called *Architecture-Driven Modernization* (ADM) (Ulrich and Newcomb, 2010) in the last years. ADM advocates carrying out re-engineering process following the principles of model-driven development: taking into account all involved artefacts as models and implementing

transformations between them.

The increasing cost of maintaining legacy systems along with the need to preserve business knowledge has turned the modernization of legacy systems into an important research field (Lewis et al., 2010). ADM provides several benefits such as ROI improvement to existing information systems, reducing development and maintenance cost and extending the life cycle of the legacy systems.

This paper proposes the elicitation of database schemas, a reverse engineering technique that follows the model-driven principles to recover a minimal relational schema from the queries embedded in the legacy source code. The reverse engineering technique to elicit relational database schema consists of two key stages:

(i) *Static analysis of source code*, which examines the Legacy source code and looks for SQL (Structure Query Language) statements embedded in the code. In this task, a model of SQL sentences is built according to a metamodel of DML (Data Manipulation Language) of SQL.

(ii) *Model transformation*, which obtains a model of the relational schema from the previous SQL sentences model. This transformation is based on a set of rules that elicits the minimal schema of the underlying database.

The objective of this proposal is to discover the minimal subset of relational elements of the legacy

database in order to improve the database to use it in the ADM process. Thus, the proposed technique rebuilds the database schema and removes the dead parts based on the embedded SQL sentences. This mechanism eradicates duplicated or unused tables, removes unused columns, and discovers new referential constraints. In this way, the software of legacy systems does not evolve in an isolated manner, but takes into account its relational databases.

The remainder of this paper is organized as follows. Section 2 summarizes related works in the database reengineering field as well as the state-of-the-art about ADM. Section 3 explains the proposed schema induction technique in detail. Section 4 presents a case study to validate the proposal. Finally, Section 5 addresses the conclusions of this work.

2 STATE-OF-THE-ART

2.1 Related Work

Research about re-engineering on applications and databases jointly is usually addressed in literature. *Reus* (Reus et al., 2006) presents a reverse engineering process based on MDA (Model-Driven Architecture) for recovering UML (Unified Modeling Language) models from PL/SQL code. *Fong* (Fong, 1997) and *Ramanathan et al.* (Ramanathan and Hodges, 1997) transform the relational models into object-oriented (OO) models to integrate them with OO applications. Also, *Cohen et al.* (Cohen and Feldman, 2003) convert parts of the application logic from the procedural style of the legacy systems to the declarative style of SQL in order to integrate them with relational databases. *Hainaut et al.* (Hainaut et al., 1996) proposes a reverse engineering process for recovering the design of relational databases and *Wu et al.* (Wu et al., 2008) discover topical structures of relational databases. Moreover, *Pérez-Castillo et al.* (Pérez-Castillo et al., 2009) propose a wrapping technique to extract Web services from relational databases that manage the data access. Finally, *Polo et al.* (Polo et al., 2007) have studied building database-driven applications.

However, related works do not address some key challenges for modernizing relational databases: (1) those works do not follow a model-driven approach in all phases of the reengineering process and (2) recovered database knowledge is not managed in an integrated and standardized manner. ADM is a

potential solution for dealing with the first challenge while KDM (Knowledge Discovery Metamodel) enables optimal knowledge management for relational databases within the ADM processes, the second challenge.

2.2 Standards Involved: ADM, KDM and QVT

The reengineering processes and MDA principles converge in ADM, an OMG standard for modernizing legacy systems (Khusidman and Ulrich, 2007). ADM is the concept of modernizing existing systems with a focus on all aspects of the current systems architecture and the ability to transform current architectures to target architectures (Pérez-Castillo et al., 2011).

The *ADM Task Force* in OMG led to several standards (OMG, 2009). The cornerstone of this set of standards is KDM (*Knowledge Discovery Meta*). KDM allows standardized representation of knowledge extracted from legacy systems by means of reverse engineering. In addition, KDM has been recently recognized as the standard ISO 19506 (ISO/IEC, 2009). The KDM metamodel provides a common repository structure that makes it possible to interchange information about software artefacts in legacy systems. KDM makes it possible to represent the PIM models in the horseshoe model. KDM can be compared with the UML standard: while UML is used to generate new code in a *top-down* manner, the ADM processes that involving KDM starts from the existing code and builds a higher level model in a *bottom-up* manner (Moyer, 2009).

The KDM metamodel (ISO/IEC, 2009) is divided into four layers (each one based on a previous layer) representing both physical and logical software assets of information systems at several abstraction levels. This work focuses on (i) *program element layer*, which provide a language-independent intermediate representation for various constructs determined by common programming languages; and (ii) *runtime resource layer*, which enables representation of high-value knowledge about legacy systems such as databases.

3 SCHEMA ELICITATION

The proposed ADM process is organized into three stages. The first stage of the modernization process aims to obtain, through reverse engineering, a set of PSM models representing each software artefact of

the legacy system. This task involves using a specific metamodel for each artefact. After that, in the second stage, a KDM model is built (using model transformations) from the PSM models recovered from the legacy system. In this case, the KDM model plays the role of a PIM model. Therefore, this model abstracts any technology-specific aspect of the legacy system. It should be borne in mind that obtaining KDM models does not end in the restructuring stage, because it is possible to restructure the KDM model itself. For example, transformations can be tailored between KDM layers. Finally, the forward engineering stage accomplishes building new and improved information systems. This stage involves PSM models to represent specific aspects related to the technological nature of each target system. In this way, the modernization process is completed according to the horseshoe model.

The aim of this modernization process is to modernize legacy systems focused on code and database as exclusive software artefacts. These artefacts undoubtedly determine three KDM models that must be obtained in the reverse engineering stage of this modernization process: (i) *KDM Inventory Model* is based on the *Source Package* of KDM. It enumerates physical artefacts of the legacy system and defines the mechanism of traceability links between all the KDM elements and their original representation in the legacy source code. (ii) *KDM Code Model* supports both *Code Package* and *Action Package*. This model aims to represent program elements and their associations at the implementation level. It includes elements supported by several programming languages such as sentences, operators, conditions, associations, control and data flows. (iii) *KDM Data Model* represents data manipulation in legacy systems. *Data Model* is based on *Data Package* and uses the foundations provided by *Code Model* related to the representation of simple data types. Also, this model can depict the relational databases used by the legacy system.

In addition to these models, the schema elicitation technique involves other models in the reverse engineering stage of this ADM process (see the shaded part of Figure 1). The database schema is elicited from the SQL embedded in the source code by means of the proposed technique, and thus it generates an *SQL Sentences Model* by means of the static analysis of legacy source code. The static analysis activity also produces the *Inventory Model* and *Code Model*. After that, the *Database Schema Model*, a model that represents the minimal schema

of the database, is obtained through the model transformation from the *SQL Sentences Model*. Finally, the needed *KDM Data Model* is obtained from the *Database Schema Model*.

At this point, both the source code and the database are represented according to KDM. Therefore the restructuring and forward engineering stages can be carried out in order to generate the modernized version of the legacy systems (see Figure 1).

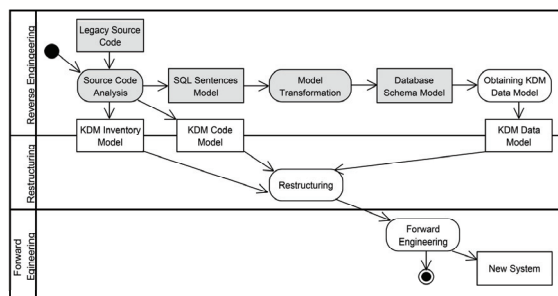


Figure 1: Schema elicitation technique based on ADM.

In order to obtain the *SQL Sentences Model* the technique analyses the legacy source code for embedded SQL sentences by means of a parser. This parser is a syntactical analyser that exhaustively scans source code. When the parser finds an SQL sentence, it translates that sentence into a model according to a metamodel of the DML (Data Manipulation Language) of SQL-92 that has been developed.

The metamodel modeling the syntax of the SQL-92 DML (ISO/IEC, 1992). It can represent the SQL operations such as *Insert*, *Select*, *Update* and *Delete* together with search conditions.

After obtaining the *SQL Sentences Model* through static analysis, the *Database Schema Model* must be obtained by mean of a model transformation. These models of relational database schemas are represented through a metamodel according to the SQL-92 standard (ISO/IEC, 1992). Deductions of the minimal database schema are based on a set of rules developed specifically for this purpose. These rules recover only a subset of the database schema elements that are handled by the SQL sentences embedded in the source code.

Rule 1. The tables that appear in any SQL sentence (Insert, Select, Update or Delete) as either source or target clauses (From, Set, Into, and so on) are created as tables in an induced database scheme.

Rule 2. The columns that are selected, added, deleted or updated in the SQL sentences are created in the corresponding tables. These tables have

previously been created through the application of Rule 1. **Rule 3.** The columns depicted through the table alias in sentences (As clause) are created in the table related to this alias. This table was created previously by means of Rule 1. **Rule 4.** The data type associated with each column can be deduced through the kind of expressions where these columns appear. For example, Like expressions → ‘string’ data type; arithmetic expressions → ‘integer, decimal or numeric’ data type, and so on. **Rule 5.** The Select sentences structured in Join mode suggest potential primary keys and foreign keys according to the pattern expressed in Figure 2. While the source column(s) of join select is/are related to the column(s) within the foreign key, the target column(s) of join select is/are related to the column(s) within the primary key. **Rule 6.** After applying the previous rules, it is possible that some tables are created without a primary key. In this case, a new column is attached to these tables as its primary key. This column is a sequential number that is generated automatically.

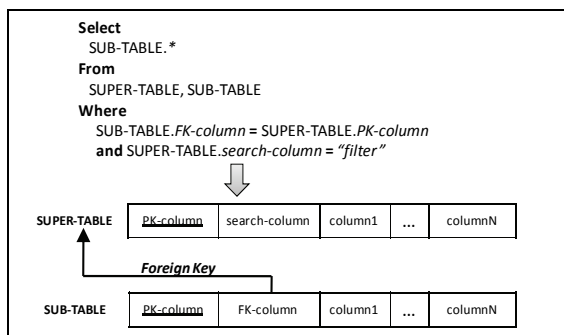


Figure 2: Pattern “join select to foreign key”.

4 CASE STUDY

The case study addresses a modernization project that is currently being carried out. The subject legacy system of this project is the intranet of the *Computer Science Faculty* at the *University of Castilla-La Mancha*. This intranet was developed five years ago by several people.

The structure of the intranet consists of five modules: (i) the *Main* module is the major module of the intranet and has the standard functionalities; (ii) the *Administration* module is in charge of administrative and office tasks; (iii) the *Old Students* module manages information related to students who were members of the faculty; (iv) the *Management* module is a module for the setup of the intranet; and

finally, (v) the *Quality* module measures and reports on the quality of the faculty.

The intranet has a typical Web architecture separated into three layers: presentation, business and persistence. The technology used to develop the presentation layer was *JSP (Java Server Pages)*, *JAVA* for business layer and *ORACLE* together with *JDBD-ODBC* for the persistence layer. The total size of this legacy system is 72.68 KLOC.

The case study establishes two research questions to analyse the *Database Schema Models* obtained through the proposal:

Q1. Are the output models complete?

Q2. Are the output models minimal with regards to the original database schema?

The question *Q1* aims to assess the completeness of the obtained database schema. A specific schema is complete when: (i) all tables in this schema model have a primary key; (ii) there are no tables without columns; and (iii) the schema model does not have any duplicated elements. Moreover, the question *Q2* takes into account the minimization of the database schema. The minimization is measured by means of the size of the obtained schema regarding the size of the source database schema. In order to measure the gain between the previous and current sizes, we use two variables: the gain related to the number of tables (1) and the gain related to the number of columns in each table (2). In these formulas, T_0 is the number of tables in the legacy database schema and $C_{0\{T_i\}}$ represents the number of columns of Table i in the legacy database. Moreover, T represents the number of tables in the improved database schema and $C_{\{T_i\}}$ is the number of columns in Table i in the obtained database schema.

$$G_T = \frac{T_0 - T}{T_0} \tag{1}$$

$$G_{C_{\{T_i\}}} = \frac{C_{0\{T_i\}} - C_{\{T_i\}}}{C_{0\{T_i\}}} \tag{2}$$

The execution of the case study was carried out by means of a tool based on the *Eclipse* platform that was developed to support the elicitation schema technique. A QVT (Queries / Views / Transformations) (OMG, 2008) model transformation is tailored in the tool from the proposed rules. The tool accomplishes several *SQL Statements Models*, a model for each source code file. Table 1 summarizes the models obtained from the legacy source code.

After that, the QVT transformations are executed through the tool using these models as input models

to obtain the output model that represents the new and minimal database schema.

Table 1: Input SQL statement models.

Module	Source Files	SQL Statements	SQL Statement Models
Main	18	23	18
Administration	8	14	8
Old Students	4	4	4
Management	1	1	1
Quality	44	60	44
Total	75	102	84

After the execution of the QVT transformations, a set of output models that depicts database segments (used for each module of the intranet) was obtained. Table 2 summarizes the results obtained for each output model. The legacy database had 140 tables and 25 tables were recovered. Table 2 shows the tables recovered for each intranet module. In addition, it presents the gain related to the tables (G_T) as well as the gain regarding the columns (G_C).

The analysis of the results obtained for these models presents several conclusions that should be considered as a response to the question $Q1$: (i) tables are usually obtained without primary keys unless Rule 6 is launch after other QVT relations; (ii) obtaining tables without columns is not usual, because any column that appears in a SQL statement is normally associated with its table.

In this case study, the QVT relations do not infer enough foreign keys, because the only QVT-implemented mechanism for inferring foreign keys is Rule 5. Indeed, the intranet source code has only two *join select* sentences due to the bad design of the legacy database.

In order to respond to the question $Q2$, the gain of the obtained database schema was also assessed. In total, 18% of the tables were recovered (25 tables) and the G_T value was 82%. With respect to the columns, the box diagram in Figure 3 shows the distributions of G_C for each intranet module. The mean per table of the G_C values was 27%, although in some modules this mean was higher. In this study, the G_C mean is lower than the G_T . However, the total gain related to the size minimization of the new database schema is significant.

5 CONCLUSIONS

This paper proposes a modernization process based on KDM. The objective of this process is the modernization of legacy source code together with legacy relational databases. For this reason, this

proposal considers two complementary sources of knowledge: (i) the schema of the legacy database, and (ii) the SQL sentences embedded in the legacy source code.

The main contribution of this paper is a mechanism named 'schema elicitation' to rebuild the database schema from the SQL sentences embedded in the source code.

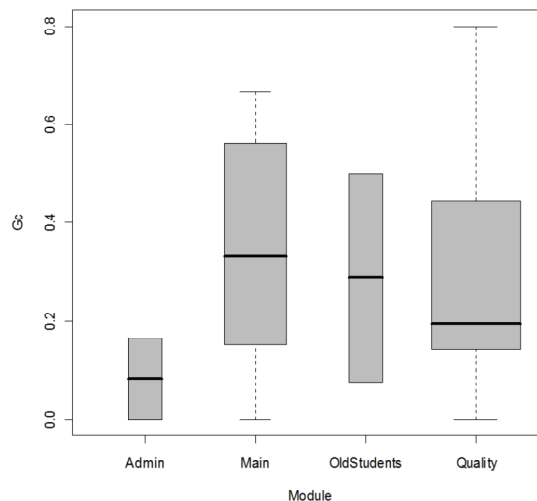


Figure 3: Box diagram of GC per module.

This mechanism removes the dead parts of the database schema such as duplicated or unused tables and unused columns. Also, this mechanism can discover new referential constraints implicit in the source code. Therefore, this proposal obtains improved and minimal database schemas used in the latter stages of the modernization process. In order to support the schema elicitation mechanism, two SQL metamodels were developed: a metamodel for representing SQL sentences embedded in source code as well as a metamodel for modelling schemas for relational databases. In addition, a set of QVT relations was tailored to transform a SQL sentences model into another database schema model.

Finally, a case study with a legacy intranet reports the main advantages of this proposal. Firstly, the obtained database schema had the adequate completeness level, thus that schema can be used as new database schema in the modernized system. Secondly, the dead parts were removed. In our case study the minimization of the size of the obtained schema was around 80% with regard to the original size.

The future extension of this research focuses on the improvement of the completeness level of the database schema models and on the detection more dead parts by means of more refined patterns. For

Table 2: The schema elements recovered through the case study and the gain of the minimal schema.

	Total Tables = 140	ALUMNOS_CEP	ANTIGUOS_ALUMNOS	ASIGNATURA	CEP	CLAVE_ALUMNOSOLD	CONFERENCIA	CONFIGURACION	CURSOS_CERO	CURSOS_CSP_POSTGR	CURSOS_TECNICAS_LET	DOCUMENTOS	EMPRESA	FECHAS_VISTAS	FORMACION_REGLOA_ASIG	GRANTIS	INFORMES_CALIDAD	MATRICULACION	NUMERO_ALUMNOS	PRESUPUESTO_INVESTIGACION	SEMINARIOS_DI	SUBCURSO_EP	TAREAS	USERS	VISTAS	Primary Keys	Foreign Keys	Tables	G _T			
Main	x			x			x					x	x				x						x	x	x	1	1	9	94%			
Administration															x									x		0	0	2	99%			
Old Students		x			x																					0	0	2	99%			
Management			x																							0	0	1	99%			
Quality				x		x		x	x	x		x	x	x	x	x	x	x	x	x	x	x				1	1	13	91%			
TOTAL	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	2	2	25	82%
Recovered Cols	4	12	3	9	1	6	1	4	5	6	7	5	4	14	5	6	3	6	4	4	5	6	11	3	1					135		
All Columns	4	13	15	28	2	19	2	7	6	7	9	9	6	18	6	6	8	6	4	7	6	7	12	3	3					213		
G_c	0	0,1	0,8	0,7	0,5	0,7	0,5	0,4	0,2	0,1	0,2	0,4	0,3	0,2	0,2	0	0,6	0	0	0,4	0,2	0,1	0,1	0	0,7					27% (Mean)		

this purpose, more case studies will be carried out in order to detect more information needs in the target database schema that must be obtained. In addition, the future work will address the next stages of the proposed modernization process such as the transformation from database models to KDM models, and then, the restructuring and forward engineering stages, which will use the previous knowledge.

ACKNOWLEDGEMENTS

This work has been supported by the *FPU Spanish Program*; by the R&D projects funded by *JCCM: ALTAMIRA (PII2I09-0106-2463)*, *INGENIO (PAC08-0154-9262)* and *PRALIN (PAC08-0121-1374)*.

REFERENCES

Bianchi, A., Caivano, D., Marengo, V. and Visaggio, G. 2003. Iterative Reengineering of Legacy Systems. *IEEE Trans. Softw. Eng.*, 29, 225-241.

Blaha, M. Year. A Retrospective On Industrial Database Reverse Engineering Projects-Part 1. In: *Proceedings of The 8th Working Conference on Reverse Engineering (WCRE '01)*, 2001 Stuttgart, Germany. Ieee Computer Society, 136-147.

Cohen, Y. and Feldman, Y. A. 2003. Automatic High-Quality Reengineering of Database Programs By Abstraction, Transformation and Reimplementation. *ACM Trans. Softw. Eng. Methodol.*, 12, 285-316.

Fong, J., 1997. Converting Relational to Object-oriented Databases. *ACM SIGMOD Record*, 26, 53-58.

Hainaut, J.-L., Henrard, J., Hick, J.-M., Roland, D. and Englebert, V. 1996. Database Design Recovery.

Proceedings of The 8th International Conference on Advances Information System Engineering. Springer-Verlag.

ISO/IEC. 1992. ISO/IEC 9075:1992, Database Language Sql.

ISO/IEC. 2009. ISO/IEC Dis 19506. *Knowledge Discovery Meta-Model (KDM)*, V1.1 (Architecture-Driven Modernization).

Khusidman, V. And Ulrich, W. 2007. Architecture-Driven Modernization: Transforming the Enterprise. Draft V.5. and [Http://www.Omg.Org/Docs/Admtf/07-12-01.Pdf](http://www.omg.org/docs/admtf/07-12-01.pdf). Omg.

Lewis, G. A., Smith, D. B. and Kontogiannis, K., 2010. A Research Agenda for Service-Oriented Architecture (Soa): Maintenance and Evolution of Service-oriented Systems. *Software Engineering Institute*.

Moyer, B., 2009. Software Archeology. Modernizing Old Systems. *Embedded Technology Journal* [Online], 1. Available: [Http://Adm.Omg.Org/Docs/Software_ArcheoLogy_4-Mar-2009.Pdf](http://Adm.Omg.Org/Docs/Software_ArcheoLogy_4-Mar-2009.Pdf) [Accessed Junio 2009].

Omg 2008. Qvt. Meta Object Facility (Mof) 2.0 Query/View/Transformation Specification. [Http://www.Omg.Org/Spec/Qvt/1.0/Pdf](http://www.omg.org/Spec/Qvt/1.0/Pdf). Omg.

Omg. 2009. Architecture-Driven Modernization Standards Roadmap. [Online]. Omg. Available: [Http://Adm.Omg.Org/Admtf%20roadmap.Pdf](http://Adm.Omg.Org/Admtf%20roadmap.Pdf) [Accessed 29/10/2009].

Pérez-Castillo, R., García-Rodríguez De Guzmán, I., Caballero, I., Polo, M. And Piattini, M. 2009. Preciso: A Reengineering Process and a Tool for Database Modernisation Through Web Services *24th ACM Symposium on Applied Computing*. P. 2126-2133.

Pérez-Castillo, R., García Rodríguez De Guzmán, I. and Piattini, M. 2011. Architecture-Driven Modernization. In: Dogru, A. H. And Bier, V. (Eds.) *Modern Software Engineering Concepts And Practices: Advanced Approaches*. Igi Global.

Polo, M., García-Rodríguez De Guzmán, I. And Piattini, M. 2007. An Mda-based Approach For Database Re-Engineering. *J. Softw. Maint. Evol.*, 19, 383-417.

Ramanathan, S. And Hodges, J. 1997. Extraction of

- Object-oriented Structures From Existing Relational Databases. *ACM SIGMOD Record*, 26, 59–64.
- Reus, T., Geers, H. And Deursen, A. V. Year. Harvesting Software for MDA-based Recovering. In: European Conference on Model Driven Architecture - Foundations and Applications, 2006 Bilbao (Spain). Springer-Verlag Berlin Heidelberg.
- Sneed, H. M., 2005. Estimating The Costs of a Reengineering Project, *IEEE Computer Society*.
- Ulrich, W. M. and Newcomb, P. H., 2010. Information Systems Transformation. Architecture Driven Modernization Case Studies, Burlington, MA, Morgan Kauffman.
- Visaggio, G., 2001. Ageing of a Data-intensive Legacy System: Symptoms and Remedies. *Journal Of Software Maintenance*, 13, 281-308.
- Wu, W., Reinwald, B., Sismanis, Y. and Manjrekar, R. Year. Discovering Topical Structures of Databases. In: *ACM SIGMOD International Conference on Management of Data*, 2008 Vancouver, Canada. Acm, 1019-1030.