BRIEF CONTENTS

INVITED SPEAKERS ................................................................. IV
ORGANIZING AND STEERING COMMITTEES .............................. V
SENIOR PROGRAM COMMITTEE .............................................. VI
PROGRAM COMMITTEE ............................................................. VII
AUXILIARY REVIEWERS ......................................................... XII
FOREWORD ............................................................... XIII
CONTENTS ................................................................. XV
INVITED SPEAKERS

Peter Géczy
AIST
Japan

Masao Johannes Matsumoto
Kyushu Sangyo University
Japan

Michele Missikoff
IASI-CNR
Italy

Barbara Pernici
Politecnico di Milano
Italy

Jianchang Mao
Yahoo! Labs
U.S.A.

Ernesto Damiani
University of Milan
Italy

Michael Papazoglou
University of Tilburg
The Netherlands
ORGANIZING AND STEERING COMMITTEES

CONFERENCE CHAIR
Joaquim Filipe, Polytechnic Institute of Setúbal / INSTICC, Portugal

PROGRAM CHAIR
José Cordeiro, Polytechnic Institute of Setúbal / INSTICC, Portugal

PROCEEDINGS PRODUCTION
Sérgio Brissos, INSTICC, Portugal
Marina Carvalho, INSTICC, Portugal
Helder Coelhas, INSTICC, Portugal
Vera Coelho, INSTICC, Portugal
Andreia Costa, INSTICC, Portugal
Bruno Encarnação, INSTICC, Portugal
Bárbara Lima, INSTICC, Portugal
Raquel Martins, INSTICC, Portugal
Carla Mota, INSTICC, Portugal
Vitor Pedrosa, INSTICC, Portugal
Vera Rosário, INSTICC, Portugal
José Varela, INSTICC, Portugal

CD-ROM PRODUCTION
Elton Mendes, INSTICC, Portugal
Pedro Varela, INSTICC, Portugal

GRAPHICS PRODUCTION AND WEBDESIGNER
Marina Carvalho, INSTICC, Portugal

SECRETARIAT AND WEBMASTER
Vitor Pedrosa, INSTICC, Portugal
Senior Program Committee

Senén Barro, University of Santiago de Compostela, Spain
Jean Bézivin, INRIA & EMN, France
Enrique Bonsón, Universidad de Huelva, Spain
Albert Cheng, University of Houston, U.S.A.
Bernard Coulette, University of Toulouse 2 - IRIT Laboratory, France
Jan Dietz, Delft University of Technology, The Netherlands
Virginia Dignum, Utrecht University, The Netherlands
Schahram Dustdar, Technical University of Vienna, Austria
António Figueiredo, University of Coimbra, Portugal
Nuno Guimarães, Lasige/Faculty of Sciences, University of Lisbon, Portugal
Dimitris Karagiannis, University of Vienna, Austria
Michel Leonard, 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
Kalle Lyytinen, Case Western Reserve University, U.S.A.
Yannis Manolopoulos, Aristotle University, Greece
José Legatheaux Martins, FCT/UNL, Portugal
Masao Johannes Matsumoto, Kyushu Sangyo University, Japan
Marcin Paprzycki, Polish Academy of Science, Poland
Alain Pirotte, University of Louvain, Belgium
Klaus Pohl, University of Duisburg-Essen, Germany
Matthias Rauterberg, Eindhoven University of Technology, The Netherlands
Coelette Rolland, Université Paris I Panthéon-Sorbonne, France
Narcyz Roztocki, State University of New York at New Paltz, U.S.A.
Abdel-badeeh Salem, Ain Shams University, Egypt
Bernadette Sharp, Staffordshire University, U.K.
Timothy K. Shih, Tamkang University, Taiwan
Alexander Smirnov, Spiiras, Russian Federation
Ronald Stamper, Measur Ltd, U.K.
Antonio Vallecillo, Universidad de Málaga, Spain
François Vernadat, European Court of Auditors, France
Merrill Warkentin, Mississippi State University, U.S.A.
PROGRAM COMMITTEE

Lena Aggestam, University of Skövde, Sweden
Miguel Angel Martinez Aguilar, University of Murcia, Spain
Patrick Albers, ESEO - Ecole Superieure D’Electronique de L’Ouest, France
Vasco Amaral, FCT/UNL, Portugal
Yacine Amirat, University Paris 12, France
Andreas Andreou, University of Cyprus, Cyprus
Colin Anthony, University of Nottingham, U.K.
Gustavo Arroyo-figueroa, Electric Research Institute of Mexico, Mexico
Wudhichai Assawinchaichote, King Mongkut’s University of Technology Thonburi, Thailand
Juan Carlos Augusto, University at Jordanstown, U.K.
Anjali Awasthi, Concordia University, Canada
Nick Bassiliades, Aristotle University of Thessaloniki, Greece
Cecilia Baranauskas, State University of Campinas - Unicamp, Brazil
Steve Barker, King’s College London University, U.K.
Balbir Barn, Middlesex University, U.K.
Lamia Hadrich Belguith, Faculté des Sciences Economiques et de Gestion, Tunisia
Nadia Bellalem, Loria, France
Orlando Belo, University of Minho, Portugal
Manuel F. Bertoa, Universidad de Málaga, Spain
Minal Bhise, Dhirubhai Ambani Institute of ICT, Gandhinagar, India
Oliver Bittel, HTWG Konstanz - University of Applied Sciences, Germany
Danielle Boulanger, Université Jean Moulin Lyon 3, France
Jean-Louis Boulanger, CERTIFER, France
Stphane Bressan, National University of Singapore, Singapore
Miguel Calejo, Declarativa, Portugal
Coral Calero, University of Castilla - La Mancha, Spain
Olivier Camp, ESEO, France
Gerardo Canfora, University of Sannio, Italy
Angélica Caro, University of Bio-Bio, Chile
Nunzio Casalino, CeRSI - LUISS University, Italy
Sergio de Cesare, Brunel University, U.K.
Maiga Chang, Athabasca University, Canada
Laurent Chapelier, Fortis Banque Luxembourg, France
Cindy Chen, University of Massachusetts Lowell, U.S.A.
Jinjun Chen, Swinburne University of Technology, Australia
Daniela Barreiro Claro, Universidade Federal da Bahia (UFBA), Brazil
Francesco Colace, Università Degli Studi di Salerno, Italy
Cesar Collazos, Universidad del Cauca, Colombia
Jose Eduardo Corcoles, Castilla-La Mancha University, Spain
Antonio Corral, University of Almeria, Spain
António Dourado Correia, University of Coimbra, Portugal
Sharon Cox, Birmingham City University, U.K.
Alfredo Cuzzocrea, University of Calabria, Italy
Jacob Cybulski, Deakin University, Australia
Mohamed Dahchour, Institut National des Postes et Télécommunications (INPT), Morocco
Suash Deb, C. V. Raman College of Engineering, India
Vincenzo Deufemia, Università di Salerno, Italy
Rajiv Dharaskar, GH Raisoni College of Engineering, India
Kamil Dimililer, Near East University, Turkey
José Javier Dolado, University of the Basque Country, Spain
PROGRAM COMMITTEE (cont.)

Juan C. Dueñas, Universidad Politécnica de Madrid, Spain
Barry Eaglestone, The University of Sheffield, U.K.
Hans-Dieter Ehrich, Technische Universitaet Braunschweig, Germany
Jean-Max Estay, Université Catholique de L’Ouest (UCO), France
Yaniv Eytani, University of Illinois at Urbana-Champaign, U.S.A.
João Faria, FEUP - Faculty of Engineering of the University of Porto, Portugal
Antonio Fariña, University of A Coruña, Spain
Antonio Fernández-caballero, Universidad de Castilla-la Mancha, Spain
Edilson Ferneda, Catholic University of Brasília, Brazil
María Joaquina Silva Costa Ferreira, Universidade do Porto, Portugal
Paulo Ferreira, INESC-ID / IST, Portugal
Filomena Ferrucci, Università Di Salerno, Italy
Mariagrazia Fugini, Politecnico di Milano, Italy
Jose A. Gallud, University of Castilla-la Mancha, Spain
Juan Garbajosa, Universidad Politécnica de Madrid - Technical University of Madrid, Spain
Leonardo Garrido, Monterrey Institute of Technology, Mexico
Peter Géczy, AIST, Japan
Joseph Giampapa, Carnegie Mellon University, U.S.A.
Paolo Giorgini, University of Trento, Italy
Raúl Giráldez, Pablo de Olavide University of Seville, Spain
Pascual González, Universidad de Castilla-la Mancha, Spain
Gustavo Gonzalez-Sanchez, Mediapro R&D, Spain
Robert Goodwin, Flinders University of South Australia, Australia
Jaap Gordijn, VU, The Netherlands
Silvia Gordillo, Universidad Nacional de la Plata, Argentina
Feliz Gouveia, University Fernando Pessoa / Cerem, Portugal, Portugal
Luis Borges Gouveia, Universidade Fernando Pessoa, Portugal
Janis Grabis, Riga Technical University, Latvia
Maria Carmen Penadés Gramaje, Technical University of Valencia, Spain
Sven Groppe, University of Lübeck, Germany
Rune Gustavsson, Blekinge Institute of Technology, Sweden
Maki K. Habib, Saga University, Japan
Abdelwahab Hamou-Ihadj, Concordia University, Canada
Christian Heinlein, Aalen University, Germany
Ajantha Herath, Richard Stockton State College of New Jersey, U.S.A.
Suvineetha Herath, Richard Stockton State College of New Jersey, U.S.A.
Francisco Herrera, University of Granada, Spain
Peter Higgins, Swinburne University of Technology, Australia
Władysław Homenda, Warsaw University of Technology, Poland
Wei-Chiang Hong, Oriental Institute of Technology, Taiwan
Jiankun Hu, RMIT University, Australia
François Jacquenet, University of Saint-Étienne, France
Ivan Jelinek, Czech Technical University in Prague, Czech Republic
Paul Johannesson, Royal Institute of Technology, Sweden
Michail Kalogiannakis, University of Crete, France
Nikos Karacapilidis, University of Patras, Greece
Nikitas Karanikolas, Technological Educational Institute of Athens (TEI-A), Greece
Stamatis Karnouskos, SAP Research, Germany
Hiroyuki Kawano, Faculty of Mathematical Sciences and Information Engineering, Nanzan University, Japan
Seungjoo Kim, Sungkyunkwan University, Korea, Republic of
Marite Kirikova, Riga Technical University, Latvia
Alexander Knapp, Universit"at Augsburg, Germany
John Krogstie, NTNU, Norway
Stan Kurkovsky, Central Connecticut State University, U.S.A.
Rob Kusters, Eindhoven University of Technology & Open University of the Netherlands, The Netherlands
Alain Leger, France Telecom Orange Labs, France
Kauko Leivisk"a, University of Oulu, Finland
Daniel Lemire, UQAM - University of Quebec at Montreal, Canada
Joerg Leukel, University of Hohenheim, Germany
Hareton Leung, Hong Kong Polytechnic University, China
Qianhui Liang, Singapore Management University, Singapore
Therese Libourel, Lirmm, France
Luis Jiménez Linares, University of de Castilla-La Mancha, Spain
Panos Linos, Butler University, U.S.A.
Gabriel Pereira Lopes, FCT/UNL, Portugal
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
Víctor López-jaquero, University of Castilla-La Mancha, Spain
Miguel R. Luaces, Universidade da Coruña, Spain
Christof Lutteroth, University of Auckland, New Zealand
Mark Lycett, Brunel University, U.K.
Cristiano Maciel, Universidade Federal de Mato Grosso, Brazil
Edmundo Madeira, Unicamp - Universidade Estadual de Campinas, Brazil
Nuno Mamede, INESC-ID, Portugal
Pierre Maret, Université de Saint Etienne, France
Herve Martin, Grenoble University, France
Katsuhisa Maruyama, Ritsumeikan University, Japan
David Martins de Matos, L2F / INESC-ID Lisboa / Instituto Superior Técnico, Portugal
Hamid Mcheick, University of Quebec at Chicoutimi, Canada
Subhas Misra, Harvard University, USDA, U.S.A.
Michele Missikoff, IASI-CNR, Italy
Ghodrat Moghadampour, Vaasa University of Technology, Finland
Pascal Molli, Loria, Université Henri Poincaré, Nancy 1, France
Francisco Montero, University of Castilla-la Mancha, Spain
Carlos León de Mora, University of Seville, Spain
Paula Morais, Universidade Portucalense, Portugal
Fernando Moreira, Universidade Portucalense, Portugal
Nathalie Moreno, University of Malaga, Spain
Haralambos Mouratidis, University of East London, U.K.
Pietro Murano, University of Salford, U.K.
Tomoharu Nakashima, Osaka Prefecture University, Japan
Paolo Napoletano, University of Salerno, Italy
Rabia Nessah, IESEG School of Management, France
Engelbert Mephu Nguifo, University Blaise Pascal - LIMOS - CNRS 6158, France
Ana Neves, knowman - Consultadoria em Gestão, Lda, Portugal
Patrick ONeil, UMass/Boston, U.S.A.
Hichem Omrani, CEPS/INSTEAD, Luxembourg
Peter Oriogun, London Metropolitan University, U.K.
Claus Pahl, Dublin City University, Ireland
José R. Paramá, University of A Coruña, Spain
Eric Pardede, La Trobe University, Australia
Rodrigo Paredes, University of Chile, Chile
Massimiliano Di Penta, University of Sannio, Italy
Laurent Péridy, IMA-UCO, France
Dana Petcu, Western University of Timisoara, Romania
Leif Peterson, TMHRI, U.S.A.
Ángeles S. Places, University of A Coruña, Spain
Geert Poels, Ghent University, Belgium
José Ragot, Centre de Recherche en Automatique de Nancy, France
Abdul Razak Rahmat, University Utara Malaysia, Malaysia
Jolita Ralyte, University of Geneva, Switzerland
Srini Ramaswamy, University of Arkansas at Little Rock, U.S.A.
Marek Reformat, University of Alberta, Canada
Hajo A. Reijers, Eindhoven University of Technology, The Netherlands
Ulrich Reimer, University of Applied Sciences St. Gallen, Switzerland
Marinette Revenu, Greyc Ensicaen, France
Nuno de Magalhães Ribeiro, Universidade Fernando Pessoa, Portugal
Simon Richir, Arts et Metiers ParisTech, France
David Rivreau, Université Catholique de L’ouest, France
Alfonso Rodriguez, University of Bio-Bio, Chile
Daniel Rodriguez, University of Alcalá, Spain
Pilar Rodriguez, Universidade Autónoma de Madrid, Spain
Oscar M. Rodriguez-Elias, UNISON, Mexico
Jose Raul Romero, University of Cordoba, Spain
Francisco Ruiz, Universidad de Castilla-La Mancha, Spain
Danguole Rutkauskiene, Kaunas University of Technology, Lithuania
Ozgur Koray Sahingoz, Turkish Air Force Academy, Turkey
Priti Srinivas Sajja, Sardar Patell University, India
Belen Vela Sanchez, Rey Juan Carlos University, Spain
Daniel Schang, ESEO, France
Sissel Guttormsen Schär, Institute For Medical Education, Switzerland
Isabel Seruca, Universidade Portucalense, Portugal
Hala Skaf-molli, INRIA Lorraine - University Henri Ponicaré, France
Pedro Soto-Acosta, University of Murcia, Spain
Chantal Soule-dupuy, Universite Toulouse 1, France
José Neuman de Souza, Universidade Federal do Ceará, Brazil
Marco Spruit, Utrecht University, The Netherlands
Hatem Ben Sta, Tunisia University, Tunisia
Martin Stanton, Manchester Metropolitan University, U.K.
Janis Stirna, Royal Institute of Technology, Sweden
Renate Strazdina, Riga Technical University, Latvia
Stefan Strecker, University of Duisburg-Essen, Germany
Chun-Yi Su, Concordia University, Canada
Ramayah T., Universiti Sains Malaysia, Malaysia
Ryszard Tadeusiewicz, Agh University of Science and Technology, Poland
Vladimir Tarasov, Jönköping University, Sweden
Sotirios Terzis, University of Strathclyde, U.K.
Claudine Toffolon, Université du Maine, France
Grigorios Tsoumakas, Aristotle University of Thessaloniki, Greece
Theodoros Tzouramanis, University of the Aegean, Greece
Athina Vakali, Aristotle University, Greece
José Ângelo Braga de Vasconcelos, Universidade Fernando Pessoa, Portugal
Michael Vassilakopoulos, University of Central Greece, Greece
Christine Verdier, LIG - University Joseph Fourier Grenoble, France
Maria-Amparo Vila, University of Granada, Spain
Bing Wang, University of Hull, U.K.

Hans Weghorn, Ba-University of Cooperative Education, Germany
Gerhard Weiss, SCCH, Austria
Graham Winstanley, University of Brighton, U.K.
Wita Wojtkowski, Boise State University, U.S.A.
Viacheslav Wolfengagen, Institute JurInfoR, Russian Federation
Robert Wrembel, Poznan University of Technology, Poland
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
Lin Zongkai, Chinese Academy of Sciences, China
**Auxiliary Reviewers**

**Michael Affenzeller**, Upper Austria University of Applied Sciences, Austria

**Rossana Andrade**, Federal University of Ceará, Brazil

**Evandro Bacarin**, UEL, Brazil

**José María Cavero Barca**, Universidad Rey Juan Carlos, Spain

**Bartosz Bebel**, Poznan University of Technology, Poland

**Ismael Caballero**, Universidad de Castilla-La Mancha (UCLM), Spain

**Jesus R. Campañá**, University of Granada, Spain

**Miguel Franklin de Castro**, Federal University of Ceará, Brazil

**Ana Cerdeira-Pena**, University of A Coruña, Spain

**Fabio Clarizia**, University of Salerno, Italy

**Fernando William Cruz**, Catholic University of Brasilia, Brazil

**Andrea Delgado**, University of the Republic, Uruguay

**Fausto Fasano**, University of Salerno, Italy

**Susana Ladra Gonzalez**, University of A Coruña, Spain

**Anastasios Gounaris**, Aristotle University of Thessaloniki, Greece

**Carmine Gravino**, University of Salerno, Italy

**Tarek Hamrouni**, Faculty of Sciences of Tunis, Tunisia

**Nantia Iakovidou**, Aristotle University, Greece

**Ioannis Katakis**, Aristotle University, Greece

**Maria Kontaki**, Aristotle University, Greece

**Eduardo Rodríguez López**, Universidad de Coruña, Spain

**Pedro Magaña**, University of Granada, Spain

**Sergio Di Martino**, Università Degli Studi di Napoli “Federico II”, Italy

**Javier Medina**, University of Granada, Spain

**Nicolás Marín**, University of Granada, Spain

**Sergio Folgar Méndez**, En xenio, S. L., Spain

**Juan Manuel Vara Mesa**, Kybele Research Group - University Rey Juan Carlos, Spain

**Isabelle Mirbel**, I3s Laboratory, France

**Mª Ángeles Moraga**, University of Castilla-La Mancha, Spain

**Thomas Natschlaeger**, Software Competence Center Hagenberg, Austria

**Matthias Nickles**, University of Bath, U.K.

**Germana Nobrega**, Universidade Católica de Brasília, Brazil

**Rocco Oliveto**, University of Salerno, Italy

**Gerald Oster**, Loria-inria Lorraine, France

**Samia Oussena**, Thames Valley University, U.K.

**Ignazio Passero**, University of Salerno, Italy

**Oscar Pedreira**, Universidade da Coruña, Spain

**Hércules Antônio do Prado**, Universidade Católica de Brasília, Brazil

**Michele Risi**, University of Salerno, Italy

**Guilermo de Bernardo Roca**, UDC, Spain

**María Dolores Ruiz**, University of Granada, Spain

**Giuseppe Scanniello**, Università Degli Studi della Basilicata, Italy

**Diego Seco**, University of A Coruña, Spain

**Boran Sekeroglu**, Near East University, Cyprus

**Manuel Serrano**, University of Castilla - La Mancha, Spain

**Yoshiyuki Shinkawa**, Ryukoku University, Japan

**Francesco Taglino**, CNR-IASI, Italy

**Eleftherios Tiakas**, Aristotle University of Thessaloniki, Greece

**Luigi Troiano**, University of Sannio, Italy

**Athanasios Tsadiras**, Aristotle University of Thessaloniki, Greece

**Corrado Aaron Visaggio**, RCOST - University of Sannio, Italy

**Fabian Wagner**, Germany

**Stéphane Weiss**, Loria, France
FOREWORD

This volume contains the proceedings of the eleventh International Conference on Enterprise Information Systems (ICEIS 2009), organized by the Institute for Systems and Technologies of Information Control and Communication (INSTICC), in cooperation with the Association for Advancement of Artificial Intelligence (AAAI), and the Association for Computation Machinery (ACM), technically co-sponsored by the Institute of Electronics Information and Communication Engineers (IEICE), and by the Workflow Management Coalition (WFMC).

This year ICEIS was held in Milan, Italy. This conference has grown to become a major point of contact between research scientists, engineers and practitioners in the area of business applications of information systems. ICEIS 2009 had five simultaneous tracks, covering different aspects related to enterprise computing, including: “Databases and Information Systems Integration”, “Artificial Intelligence and Decision Support Systems”, “Information Systems Analysis and Specification”, “Software Agents and Internet Computing” and “Human-Computer Interaction”. Papers published in each and every track describe state-of-art research work that is often oriented towards real world applications and highlight the benefits of Information Systems and Technology for industry and services, thus making a bridge between the Academia and the Enterprise worlds.

Following the trend of previous editions, ICEIS 2009 also had a number of satellite workshops, related to the field of the conference, including the following nine international workshops: 9th Int’l Workshop on Pattern Recognition in Information Systems; 7th Int’l Workshop on Modelling, Simulation, Verification and Validation of Enterprise Information Systems; 7th Int’l Workshop on Security In Information Systems; 6th Int’l Workshop on Natural Language Processing and Cognitive Science; 3rd Int’l Workshop on RFID Technology - Concepts, Applications, Challenges, 3rd Int’l Workshop on Human Resource Information Systems, 1st Int’l Workshop on Future Trend of Model-Driven Development, 1st Int’l Workshop on Ontology for e-Technologies and the 1st Joint Workshop on Advanced Technologies and Techniques for Enterprise Information Systems.

ICEIS 2009 received 644 paper submissions from 70 countries on all continents. 81 papers were published and presented as full papers, i.e. completed work (8 pages/30’ oral presentation) and 171 papers, reflecting work-in-progress, were accepted and orally presented as short papers (6 pages/20’ oral presentation). Furthermore, 87 contributions were accepted and presented as posters.

These numbers, leading to a “full-paper” acceptance ratio of 12%, and a total oral acceptance ratio below 40%, show the intention of preserving a high quality forum for the next editions of this conference. Additionally, as usual in the ICEIS conference series, a number of invited talks, presented by internationally recognized specialists in different areas, have positively contributed to reinforce the overall quality of the Conference and to provide a deeper understanding of the Enterprise Information Systems field.
This year, the book of Selected Papers, usually published after the conference by Springer-Verlag in the “Lecture Notes in Business Information Processing” series, has been anticipated and made available at the conference, including all full papers. For this reason the ICEIS full papers are not included in the proceedings.

The program for this conference required the dedicated effort of many people. Firstly, we must thank the authors, whose research and development efforts are recorded here. Secondly, we thank the members of the program committee and the additional reviewers for a valuable help with their expert reviewing of all submitted papers. Thirdly, we thank the invited speakers for their invaluable contribution and for taking the time to synthesise and prepare their talks. Fourthly, we thank the workshop chairs whose collaboration with ICEIS was much appreciated. Finally, special thanks to all the members of the INSTICC organizing committee, especially Vitor Pedrosa, whose diligence and dedication was fundamental for the success of this conference.

Two awards will be given to the best papers presented at the conference: one for the best paper award and one for the best student paper award, mainly based on the classifications provided by the Program Committee members and the oral presentation quality.

We wish you all an exciting conference and an unforgettable stay in Milan. We hope to meet you again next year for the 12th ICEIS, details of which will be readily available at http://www.iceis.org.

Conference Chair

Joaquim Filipe, Polytechnic Institute of Setúbal / INSTICC, Portugal

Program Chair

José Cordeiro, Polytechnic Institute of Setúbal / INSTICC, Portugal
CONTENTS

INVITED SPEAKERS

KEYNOTE SPEAKERS

HUMAN BEHAVIOR AND INTERACTIONS IN WEB ENVIRONMENTS
Peter Géczy

SERVICE COMPUTING EIS, WORLD PANIC AND OUR ROLE CHANGE
Masao Johannes Matsumoto

LET’S SEMANTICISE THE WORLD!! ... OR NOT??
Michele Missikoff

DESIGN OF ADAPTIVE WEB SERVICES
Barbara Pernici

MACHINE LEARNING IN ONLINE ADVERTISING
Jianchang Mao

RISK-AWARE COLLABORATIVE PROCESSES
Ernesto Damiani

FORESIGHT & RESEARCH PRIORITIES FOR SERVICE ORIENTED COMPUTING
Michael Papazoglou

DATABASES AND INFORMATION SYSTEMS INTEGRATION

SHORT PAPERS

EXPERIENCES OF ERP USE IN SMALL ENTERPRISES
Paivi Iskanius, Raija Halonen and Matti Mottonen

BUSINESS INTELLIGENCE BASED ON A WI-FI REAL TIME POSITIONING ENGINE - A Practical Application in a Major Retail Company
Vasco Vinhas, Pedro Abreu and Pedro Mendes

DIRECTED ACYCLIC GRAPHS AND DISJOINT CHAINS
Yangjun Chen

AN OBJECT MODEL FOR THE MANAGEMENT OF DIGITAL IMAGES
S. Khaddaj and Andreas Hoppe

A MAPREDUCE FRAMEWORK FOR CHANGE PROPAGATION IN GEOGRAPHIC DATABASES
Ferdinando Di Martino, Salvatore Sessa, Giuseppe Polese and Mario Vacca

ESTABLISHING TRUST NETWORKS BASED ON DATA QUALITY CRITERIA FOR SELECTING DATA SUPPLIERS
Ricardo P. del Castillo, Ismael Caballero, Ignacio García-Rodríguez, Macario Polo, Mario Piattini and Eugenio Verbo

ALGORITHMS FOR EFFICIENT TOP-K SPATIAL PREFERENCE QUERY EXECUTION IN A HETEROGENEOUS DISTRIBUTED ENVIRONMENT
Marcin Gorawski and Kamil Dowlaszewicz

XV
Abstract: Nowadays, organizations may have Web portals tailoring several websites where a wide variety of information is integrated. These portals are typically composed of a set of Web applications and services that interchange data among them. In this setting, there is no way to find out how the quality of the interchanged data is going to evolve successively. A framework is proposed for establishing trust networks based on the Data Quality (DQ) levels of the interchanged data. We shall consider two kinds of DQ: inherent DQ and pragmatic DQ. Making a decision about the selection of the most suitable data supplier will be based on the estimation of the best expected pragmatic DQ levels. In addition, an example is presented to illustrate framework operation.

1 INTRODUCTION

Currently, companies usually have several interrelated Web portals. These Web portals integrate different Web applications. Indeed, there may be external links to Websites of other organizations. Used information may not be stored in a centralized manner in order to be shared by all applications, but each application typically manages its own data (Yin et al., 2007). There is a data flow among these Web applications. Each application, site or service in the Web portal (named node in this paper) can act as a supplier or consumer of data in any given moment. The set of participating nodes is called data networks in (Cai and Shankaranarayanan, 2007). In these networks, a business process in a node may have defined several data source nodes that are not mutually exclusive. Thus, a certain node for a certain business process is entitled to collect data from its supplier nodes. However, the node only collects required data from one of the nodes at any given moment.

A problem of Data Quality (DQ) can appear in the scenario described above: If a node of the network needs to acquire pieces of data from another node, it might not meet the quality of incoming data (Cai and Shankaranarayanan, 2007) and thus, it may use data with inadequate levels of DQ. In other words, a Web application can only understand the quality of incoming data; the so-called ‘inherent DQ’. This DQ is the degree to which data accurately reflects the real-world object that the data represents (English, 1999). In spite of the node knows its ‘inherent DQ’, it does not understand how much quality the incoming data has until it is interchanged and used; this DQ is called ‘pragmatic DQ’. This DQ is the degree of node customer satisfaction derived by the use that it is made of pieces of data (English, 1999). Impossibility to meet the pragmatic DQ in this scenario is due to two main reasons. (1) Even in an hypothetical case of a node knowing the inherent DQ of the provided data, the DQ could be different after the acquisition, since pragmatic DQ is dependent on the context (Strong et al., 1997). (2) In the case of having different suppliers for the same information need (Wu and Marian, 2007), they are expected to provide data with different expected pragmatic DQ levels.

Low levels of DQ affect the overall efficiency of the organization (Caballero et al., 2004). According to (Eppler and Helfert, 2004), the cost of preventing DQ problems is lower than the cost of detecting and repairing them. So in this scenario of Web portals interchanging data, it would be reasonable to prevent DQ problems before they appear. One way to achieve this prevention, or at least minimize its effect, can
consist of selecting the best data supplier for a task.

This paper proposes a framework based on trust networks, which can be used by a node of the network to estimate the expected pragmatic DQ. These Trust Networks allow taking into account the data provenance (Prat and Madnick, 2008), i.e. all processing history of data from its source. The goal is to select, in a heuristic manner, among all available nodes which is the one offering higher DQ levels. In each network, expected pragmatic DQ will be estimated between each pair of nodes creating different supply chains (Nicolaou and McKnight, 2006). Each of these supply chains will provide, in the end, a \( DQ \) pragmatic value that represents the data provenance of the chain. This will allow choosing the most suitable data supplier. The remainder of this paper is structured as follows: the second section reviews related work. The third section presents the proposed framework and illustrates its usage by means of an example. The final section presents the conclusions and future work.

2 RELATED WORK

Many authors agree that data has quality if it fits the intended use for which it was created (Batini and Scannapieco, 2006; Strong et al., 1997). Inadequate levels of DQ in an organizational Information System will have a negative impact on the business performance (Caballero et al., 2004). Therefore, organizations should take into account DQ issues in order to improve their performance (Al-Hakim, 2007). Due to the existence of data networks (Cai and Shankaranarayanan, 2007), assessing the DQ of each Web node in the data network is not enough (Caro et al., 2008; Eppler et al., 2003). One of the most interesting strategies for tackling the study of DQ for data network context, is to break it down into ‘minor qualities’ known as DQ dimensions.

According to English (English, 1999), assessment of the inherent DQ, the DQ dimensions belonging to the intrinsic category given by (Strong et al., 1997), (Accuracy, objectivity, believability and reputation), may be used. On the other hand, the pragmatic DQ can be assessed through DQ dimensions of the contextual category (relevancy, added value, timeliness, completeness, amount of data) given by (Strong et al., 1997). For our proposal, we will be interested in measuring not only the inherent DQ of the pieces of data that it are interchanged between each pair of nodes, but we also hope to estimate how usable they will be for an application (Even and Shankaranarayanan, 2007). In order to estimate the \( Pragmatic\ DQ \), the objective is to assist in the selection of the optimal data supplier, using DQ as a discriminator (Al-Hakim, 2007).

Moreover, the research in the DQ field suggests moving the focus from Information Systems to \textit{Information Products (IP)} (Wang et al., 1998). This approach proposes considering pieces of information as products because standard techniques for managing DQ, like \textit{Total Data Quality Management (TDQM)} (Wang, 1998), can be applied. IP-MAP graphical notation has emerged for depicting IPs (Shankaranarayanan et al., 2000). IP-MAP indicates how an IP is created during the manufacturing process. Moreover, an IP-XML file is used for representing IP-MAP meaning through metadata that can be interchanged (Cai and Shankaranarayanan, 2007).

In order to efficiently assess the quality of data, knowledge of where pieces of data have been provided from is necessary. Moreover, in this assessment, it is essential to know the historical transport of pieces of data. According to (Simmhan et al., 2005) data provenance is “information that helps to determine the derivation history of a data product, starting from its original sources”. This approach has been used in data sharing and data integration. For instance, provenance information is used to determine data updates, to explain relationships between source and target nodes that interchange data, and so on (Buneman and Tan, 2007).

Finally, the trust networks consist of a set of transitive relations of trust between people, organizations and information systems connected in an intercommunicated environment (Yin et al., 2007). In a specific semantic context, trust is transitive and may be derived from the network (Josang et al., 2007). Usefulness of these networks is in the ability to make trust-based decisions: these networks can infer trust in nodes that are not communicated directly (Josang et al., 2007). This is a key advantage of these networks, because an application or service on a Web site can choose the provider with a greater degree of trust. In this selection, the application or site will not be aware of all providers in the supply chain that are behind it (Josang et al., 2007). The Application or site knows only the nodes directly related to it.

3 PROPOSED FRAMEWORK

The selection of a data supplier could be made, taking as a basis, the observation of \textit{inherent DQ} in each node acting as data supplier. However, the framework proposes to estimate the expected \textit{pragmatic DQ} of the pieces of data supplied by each node in the data network (Tinglong and Xiangtong, 2007) as a crite-
tion for selecting the best supplier node. Therefore, finding an approximate value that synthesizes the expected pragmatic DQ (English, 1999) along a supply route in the network is proposed.

The structure of the proposed framework is the following: the entire process for creating a trust network will be governed by a ‘trust network creation’ algorithm which uses three components that are also defined in the framework. (1) ‘Matching method’ selects a subset of nodes involved in the data network which can be candidates belonging to the trust network of a given node. (2) ‘Estimation of Expected Pragmatic DQ’ method which is responsible for estimating an approximated value of the expected pragmatic DQ along the supply chains in the trust network. (3) ‘Function of data supplier selection’ allows selecting the most appropriate data supplier in terms of expected pragmatic DQ. The following paragraphs explain the details of each component.

### 3.1 Trust Network Creation Algorithm

To define the scope of a trust network our framework incorporates an algorithm that will define the limits of network on which pragmatic DQ is estimated. It starts from the node that requires pieces of data. The algorithm establishes the nodes within the trust network that it attempts to develop. The trust network is going to be built through transitive relations. These relationships are identified by a matching process. Through the algorithm (see Algorithm 1), the network is built starting from the ‘node’ which tries to select the best data supplier for an Information Product (IP) manufacturing process (Wang, 1998). An XML-Based description of the IP-MAP diagram corresponding to the manufacturing process can be made by IP-XML (Cai and Shankaranarayanan, 2007). The IP-XML file, containing information about the data network, will be one of the arguments of the matching function. Each node will recursively ask its successive suppliers through the matching function ‘getDirectSuppliers’. The algorithm also accepts the argument ‘threshold’ as a way to stop recursion (Josang et al., 2007). This limitation tries to minimize derived problems of cycles on the network. The threshold indicates the depth achieved by the algorithm during the node search (Tinglong and Xiangtong, 2007). Once the algorithm arrives at the deepest point of the different supply routes, the estimated values of expected pragmatic DQ (estimated trust) go backward within argument ‘measures’. When the algorithm reaches back to the consumer node, the node will be in disposition to select the most suitable data supplier by means of the function ‘selectOptimal’.

#### Algorithm 1: SelectSupplier

```plaintext
input:
node: It is the consumer node where trust network will be built
suppliers: It represents IP-XML info associated with node
threshold: It is the maximum number of data interchanges

output:
supplierNode: It is the optimal node to provide data to the node

1 begin
2 if threshold = 0 then
3 supplierNode ← node.getInherentDQ()
4 else
5 end
6 if measures () = 0
7 suppliers () ← node.getDirectSuppliers(ipxml)
8 foreach sup ∈ suppliers do
9 measures ← measures ∪ selectSupplier(sup, sup.ipxml.threshold-1)
10 end
11 supplierNode ← selectOptimal(measures.getExpectedPragmaticDQ ())
12 return supplierNode
13 end
14 end
```

### 3.2 Matching Method

The matching method can determine the transitivity of trust in the network (Josang et al., 2007), i.e. the transitivity of pragmatic DQ. This method analyzes the IP-MAP diagram of each node and contrasts each IP-MAP in trying to find an overlapping point where offering fits demand (Cai and Shankaranarayanan, 2007). These overlapping points are determined through the comparison between process blocks in different IP-MAP diagrams. IP-MAP is a graphical notation to represent the elaboration process of Information Products (IP) (Shankaranarayanan et al., 2000; Wang, 1998). IP-MAP includes a set of construct blocks to depict the raw input/output data, processing, data storage, decisions and so on. For each process, the correspondence between the raw input data blocks and raw output data block in both IP-MAP diagrams is examined. This activity requires a mechanism that indicates the semantics of involved process in the data networks. Due to this semantics, the matching method will identify the overlapping points. In this paper, we propose to use IP-MAP (Cai and Shankaranarayanan, 2007). However, others mechanisms could be used for this task as Business Process Modeling Notation (BPMN) or activities diagrams. The algorithm (see Algorithm 1), through the matching method, determines the subset of trust network nodes among all data network nodes. At this moment, the algorithm is at the deepest point of recursion (see Algorithm 1), and has established the entire network of nodes involved in the assessment of trust (pragmatic DQ) through the matching method.

### 3.3 Estimating Expected Pragmatic DQ

At this stage, the framework should estimate the expected pragmatic DQ in each set of suppliers.
The pragmatic DQ will be spread backward until it reaches the basis node consumer, allowing it to select the best supplier (Eppler et al., 2003). This pragmatic DQ has to synthesize, somehow, the value of historic, pragmatic and inherent DQ that there is behind each supplier in its supply chain (Al-Hakim, 2007). These supply chains represent the data provenance of each network node. Therefore, each node on the network has an associated inherent DQ value based on the DQ of supplied data for certain processes, and another estimated pragmatic DQ value. The inherent DQ value will be measured under the following assumptions. (1) DQ dimensions must be established previously for measuring the inherent DQ (Eppler et al., 2003). These DQ dimensions are the same for each set of supplied data, and must be compatible with all network nodes. (2) It will use a synthesizing numerical value of inherent DQ for each node in the network. This value represents the degree of trust exhibited in the network (Yin et al., 2007). To obtain this unique value, a process of grouping values of the different dimensions has to be executed. It involves the following actions. (2a) Summarizing and grouping functions like averages, totals, maximums, and so on. (2b) For non-numerical dimensions, a set of linguistic labels and soft-computing techniques to obtain a numerical value. (2c) To normalize all DQ dimensions the same scale ‘S’ is used which is defined by a minimum and maximum value.

$$scale(S) = S_{\text{max}} - S_{\text{min}}$$ (1)

Each node of the trust network offers data with an expected pragmatic DQ level ($Q_P$). The estimation of this $Q_P$ value is carried out by means of the following heuristics. These are based on other similar studies as (Yin et al., 2007).

**Heuristic 1.** Pragmatic DQ of a certain node depends on both Inherent DQ of this node and Pragmatic DQ of all nodes which interchange pieces of data with the node.

**Heuristic 2.** The weighting of each Pragmatic DQ value, in each node that affect source node, is related to difference between Inherent DQ and Pragmatic DQ for each node.

Therefore, $Q_P$ value depends on its inherent DQ ($Q_I$) and on estimated pragmatic DQ of its set of suppliers. Both terms are given a node-dependent weight $\alpha$ and $\beta$ (see (5) and (6)). For taking into account the pragmatic DQ values of the suppliers, it will make an average on every $Q_P$ belong to set of suppliers ($\{\text{suppliers}\}$). The heuristic 2 is used to obtain $W_k$: the weight associated with each term $k$ belonging to $\{\text{suppliers}\}$ ($W_k$) will be proportional to how $Q_P$ and $Q_I$ differ in each node.

$$W_k = 1 - \frac{|Q_{pk} - Q_{ik}|}{\text{scale}(S)}$$ (2)

In (3) (using formula (2)), the suppliers’ $Q_P$ is summarized. This term is identified as $\sigma_{pk}$ which is based on provenance-based believability assessment presented in (Prat and Madnick, 2008):

$$\sigma_{pk} = \frac{\sum_{k\in\{\text{suppliers}\}} |W_k| \cdot Q_{pk}}{|\{\text{suppliers}\}|}$$ (3)

Taking into account (2), (3) and also the inherent DQ, the estimated value of $Q_P$ in the node $k + 1$ is as:

$$Q_{pk+1} = \alpha \cdot Q_{Ik+1} + \beta \cdot \sigma_{pk}$$ (4)

This formula is a recurrent function which allows to propagating back $Q_P$ values towards initial node. Moreover the framework establishes $\alpha$ and $\beta$ weights in (5) and (6). For a specific node, if suppliers’ $Q_P$ varies greatly, it will give more weight to the $Q_I$ of that node. In addition, there are two exceptional cases: on one hand, if the algorithm is at the network limits, and hence suppliers do not exist, it only considers $Q_I$, so $\alpha = 1$. And on the other hand, if there is only one supplier, and therefore cannot check the disparity of $Q_P$, then $\alpha = \frac{1}{2}$ for $Q_I$ and $\sigma_{pk}$ have the same weight.

$$M = \max(|Q_{Pk}| n \in \{\text{suppliers}\})$$

$$m = \min(|Q_{Pk}| n \in \{\text{suppliers}\})$$

$$\alpha = \begin{cases} 1 & \text{if } |\{\text{suppliers}\}| = 0 \\ \frac{1}{2} - \frac{m - m}{\text{scale}(S)} & \text{if } |\{\text{suppliers}\}| = 1 \\ 1 & \text{if } |\{\text{suppliers}\}| > 1 \end{cases}$$ (5) (6)

### 3.4 Function of Data Supplier Selection

At this stage, the proposed algorithm has returned all pragmatic DQ values for each origin node’s suppliers. At this point, the node will select the most suitable supplier according to the expected pragmatic DQ through a selection function (Al-Hakim, 2007; Tinglong and Xiangtong, 2007). The selection function must take into account the acquired knowledge of data provenance. This function aims to select the network node which will provide data. The selection function can implement criteria as simple as choosing the greatest $Q_P$ value among all their supply nodes. However, the selection function could be more sophisticated, and consider for example: the $Q_P$ evolution over time, combining several estimated measures, taking into account the quality/cost relationship and so on.
4 USING THE FRAMEWORK

In this section, we present an example to illustrate the use of framework. The Figure 1 depicts the data network of an organization. The algorithm creates a trust network for a certain task in a certain node. In our example, the certain task is ‘stock updating’ and the certain node is sales Web application (see Figure 1). The algorithm uses the IP-MAP diagrams during the process of matching. The sales application node obtains the IP-XML of those nodes with which it is logically interconnected (production, intranet and corporative website) (see Figure 1). The matching method has verified that two of the three, both the intranet and production nodes, can act as data suppliers for the IP in the consumer node. In this case, the matching method has contrasted that some data destinations in the IP-MAP of these nodes contain data sources in IP-MAP of the sales Web application node. The matching method is executed successively until all supply routes are established. The trust network based on DQ will be applied on the recently created network (see Figure 2).

Figure 1: Network of an organization.

Figure 2: Created Trust Network.

For the sake of estimating the pragmatic DQ, each node of the trust network established previously for the case of stock updating in sales Web application should be borne in mind. In this stage, the algorithm will start estimations of expected pragmatic DQ in different network nodes. The network (see Figure 3) details inherent DQ values, offered initially by each network node. The scale of DQ values is between 1 and 10. In addition, the Figure 3 illustrates the first $Q_p$ values (Warehouse and Assembly Line nodes). These values are propagated within the network towards the origin node (sales Web application). In this case, the absence of suppliers makes $\alpha = 1$ which implies that $Q_P = Q_r$. Then, expected pragmatic DQ of the production node is calculated based on Warehouse and Assembly Line nodes (see Figure 4). The weights are $\alpha = 0.1$ and $\beta = 0.9$ because $Q_{P_{\text{Assembly Line}}} = 5$ and $Q_{P_{\text{Warehouse}}} = 4$, whose difference is 1. Therefore $Q_{P_{\text{Production}}} = 0.1 \cdot 6 + 0.9 \cdot \frac{1 - 0.4}{2} + \frac{1 - 0.5}{2} = 4.65$. The estimated $Q_{P_{\text{Production}}}$ value is offered to intranet and sales application nodes. Nevertheless, sales Web application node disposes of this value only, hence $Q_{P_{\text{Intranet}}}$ must be also estimated (see Figure 4). Finally, expected pragmatic DQ of the intranet node is estimated (see Figure 5). The weights are $\alpha = 0.5$ and $\beta = 0.5$ because intranet node has a single supplier node; hence $Q_{P_{\text{Intranet}}} = 0.5 \cdot 7 + 0.5 \cdot \frac{(1 - 0.135)\cdot 4.65}{1} = 5.51$.

After all pragmatic DQ values have been estimated in the trust network, the optimal supply node can be selected. We must remember that in this case the selection function is as simple as selecting the greatest $Q_p$ value. In the example (see Figure 5), the sales Web application will take data for updating the stock from the intranet, because the trust ($Q_p$) of this node with 5.51 is greater than the one of the production node whose value is 4.65.

Figure 3: Trust calculations in the network (Step I).

Figure 4: Trust calculations in the network (Step II).

Figure 5: Trust calculations in the network (Step III).
5 CONCLUSIONS

This paper has proposed a framework based on trust networks applied to data networks. The framework estimates an expected value at each node in the supply chain, taking into account the remaining nodes that supply data to it. The presented framework is able to determine which data supplier offers the most suitable expected pragmatic $DQ$ in each provenance scenario. The proposed framework uses, undoubtedly, an approximated measurement, therefore there is no guarantee of finding the optimal provider in all situations. In the future, we will work on two key aspects. (1) It will be validate in empirical manner as well as by means of simulation or analytical evaluation. (2) We will provide several selection functions which take into account other factors as quality/cost relationship or historical data in order to increase support to decision-making in these networks.

ACKNOWLEDGEMENTS

This research is part of the projects ESFINGE (TIN2006-15175-C05-05), DQNet (TIN2008-04951-E) and HERMES (TSI-020100-2008-155) supported by the Spanish Ministerio de Educació n y Ciencia and project IVISCUS (PAC08-0024-5991) supported by the Consejería de Educación y Ciencia de Comunidades de Castilla - La Mancha.

REFERENCES


