

Communications in Computer and Information Science

737

Commenced Publication in 2007

Founding and Former Series Editors:

Alfredo Cuzzocrea, Dominik Ślęzak, and Xiaokang Yang

Editorial Board

Simone Diniz Junqueira Barbosa

*Pontifical Catholic University of Rio de Janeiro (PUC-Rio),
Rio de Janeiro, Brazil*

Phoebe Chen

La Trobe University, Melbourne, Australia

Xiaoyong Du

Renmin University of China, Beijing, China

Joaquim Filipe

Polytechnic Institute of Setúbal, Setúbal, Portugal

Orhun Kara

TÜBİTAK BİLGEM and Middle East Technical University, Ankara, Turkey

Igor Kotenko

*St. Petersburg Institute for Informatics and Automation of the Russian
Academy of Sciences, St. Petersburg, Russia*

Ting Liu

Harbin Institute of Technology (HIT), Harbin, China

Krishna M. Sivalingam

Indian Institute of Technology Madras, Chennai, India

Takashi Washio

Osaka University, Osaka, Japan

More information about this series at <http://www.springer.com/series/7899>

Chiara Francalanci · Markus Helfert (Eds.)

Data Management Technologies and Applications

5th International Conference, DATA 2016
Colmar, France, July 24–26, 2016
Revised Selected Papers

Editors

Chiara Francalanci
Department of Electronics and Information
Politecnico di Milano
Milan
Italy

Markus Helfert
School of Computing
Dublin City University
Dublin
Ireland

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-319-62910-0 ISBN 978-3-319-62911-7 (eBook)
DOI 10.1007/978-3-319-62911-7

Library of Congress Control Number: 2017945731

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The present book includes extended and revised versions of a set of selected papers from the 5th International Conference on Data Management Technologies and Applications (DATA 2016), held in Lisbon, Portugal, during July 24–26, 2016.

We received 50 paper submissions from 26 countries, of which 18% are included in this book. The papers were selected by the event chairs and their selection is based on a number of criteria that include the classifications and comments provided by the Program Committee members, the session chairs' assessment, and also the program chairs' global view of all papers included in the technical program. The authors of selected papers were then invited to submit a revised and extended version of their papers having at least 30% innovative material.

The 5th International Conference on Data Management Technologies and Applications (DATA) aims to bring together researchers, engineers and practitioners interested in databases, data warehousing, data mining, data management, data security and other aspects of information systems and technology involving advanced applications of data.

The papers selected to be included in this book contribute to the understanding of relevant trends of current research on business analytics, data management and quality, ontologies and the Semantic Web, and databases and data security. They substantially contribute to the literature by providing interesting use cases, demonstrating the application of modern big data design and management techniques. The interdisciplinary approach of the DATA conference is a fundamental enabler of this scientific contribution. Particularly, the ability to attract both academics and practitioners and a focus on technology application represent important drivers of the practical consequence of the papers published in this book. The papers address a range of specific topics including: data consistency on agnostic fault-tolerant systems, improving performances of an embedded relational database management system, visual citation tracing, efficient multi-domain data processing, pay-as-you-go data quality management, real-time analytics of streaming data, computation of belief combination rules, identifying conversational message threads, and ETL processes specification using a pattern-based ontology. Overall, they provide an overview of the state of the art, with a focus on data and on their impact on business processes. Organizational policies, technical choices, and economic benefits from the exploitation of modern technologies are largely addressed by the papers published in this book, with a broad set of references to the most useful literature on the subject.

We would like to thank all the authors for their contributions and also the reviewers who helped ensure the quality of this publication.

February 2017

Chiara Francalanci
Markus Helfert

Organization

Conference Chair

Markus Helfert
Dublin City University, Ireland

Program Chair

Chiara Francalanci
Politecnico di Milano, Italy

Program Committee

| | |
|--------------------------|---|
| Muhammad Abulaish | South Asian University, India |
| Hamideh Afsarmanesh | University of Amsterdam, The Netherlands |
| Markus Aleksy | ABB Corporate Research Center, Germany |
| Christos Anagnostopoulos | University of Glasgow, UK |
| Nicolas Anciaux | Inria Paris-Rocquencourt, France |
| Kenneth Anderson | University of Colorado, USA |
| Keijiro Araki | Kyushu University, Japan |
| Bernhard Bauer | University of Augsburg, Germany |
| Andreas Behrend | University of Bonn, Germany |
| Fevzi Belli | Izmir Institute of Technology, Turkey |
| Karim Benouaret | Université Claude Bernard Lyon 1, France |
| Jorge Bernardino | Polytechnic Institute of Coimbra - ISEC, Portugal |
| Francesco Buccafurri | University of Reggio Calabria, Italy |
| Dumitru Burdescu | University of Craiova, Romania |
| Kung Chen | National Chengchi University, Taiwan |
| Yangjun Chen | University of Winnipeg, Canada |
| Byron Choi | Hong Kong Baptist University, Hong Kong, SAR China |
| Christine Collet | Grenoble Institute of Technology, France |
| Agostino Cortesi | Università Ca' Foscari di Venezia, Italy |
| Theodore Dalamagas | Athena Research and Innovation Center, Greece |
| Bruno Defude | Institut Mines Telecom, France |
| Steven Demurjian | University of Connecticut, USA |
| Stefan Dessloch | Kaiserslautern University of Technology, Germany |
| Fabien Duchateau | Université Claude Bernard Lyon 1/LIRIS, France |
| Todd Eavis | Concordia University, Canada |
| Mohamed Y. Eltabakh | Worcester Polytechnic Institute, USA |
| Markus Endres | University of Augsburg, Germany |
| Sergio Firmenich | Universidad Nacional de La Plata, Argentina |
| Kehan Gao | Eastern Connecticut State University, USA |
| Roberto García | Universitat de Lleida, Spain |

| | |
|---------------------------------|--|
| Jérôme Gensel | Université Grenoble Alpes, France |
| Paola Giannini | University of Piemonte Orientale, Italy |
| Giorgos Giannopoulos | Athena Research and Innovation Center, Greece |
| J. Paul Gibson | Mines-Telecom, Telecom SudParis, France |
| Matteo Golfarelli | University of Bologna, Italy |
| Janis Grabis | Riga Technical University, Latvia |
| Jerzy Grzymala-Busse | University of Kansas, USA |
| Mena Habib | Maastricht University, The Netherlands |
| Raju Halder | Indian Institute of Technology Patna, India |
| Waqar Haque | University of Northern British Columbia, Canada |
| Andreas Henrich | University of Bamberg, Germany |
| Jose Luis Arciniegas Herrera | Universidad del Cauca, Colombia |
| Jang-Eui Hong | Chungbuk National University, Korea, Republic of |
| Tsan-Sheng Hsu | Institute of Information Science, Academia Sinica, Taiwan |
| Ivan Ivanov | SUNY Empire State College, USA |
| Wang Jianmin | Tsinghua University, China |
| Konstantinos Kalpakis | University of Maryland Baltimore County, USA |
| Dimitris Karagiannis | University of Vienna, Austria |
| Maurice van Keulen | University of Twente, The Netherlands |
| Benjamin Klöpper | ABB Corporate Research, Germany |
| Mieczyslaw Kokar | Northeastern University, USA |
| Kostas Kolomvatsos | University of Thessaly, Greece |
| John Krogstie | NTNU, Norway |
| Martin Krulis | Charles University, Czech Republic |
| Konstantin Läufer | Loyola University Chicago, USA |
| Dominique Laurent | ETIS Laboratory CNRS UMR 8051, Cergy-Pontoise University, ENSEA, France |
| Sangkyun Lee | TU Dortmund, Germany |
| Raimondas Lencevicius | Nuance Communications, USA |
| Haikun Liu | Huazhong University of Science and Technology, China |
| Ricardo J. Machado | Universidade do Minho, Portugal |
| Zaki Malik | Wayne State University, USA |
| Keith Marsolo | Cincinnati Children's Hospital Medical Center, USA |
| Miguel A. Martínez-Prieto | University of Valladolid, Spain |
| Florent Masseglia | Inria, France |
| Fabio Mercorio | University of Milano-Bicocca, Italy |
| Dimitris Mitrakos | Aristotle University of Thessaloniki, Greece |
| Stefano Montanelli | Università degli Studi di Milano, Italy |
| Bongki Moon | Seoul National University, Korea, Republic of |
| Gianluca Moro | Università di Bologna, Italy |
| Mirella M. Moro | Federal University of Minas Gerais (UFMG), Brazil |
| Mikhail Moshkov | KAUST, Saudi Arabia |
| Josiane Mothe | Université de Toulouse, France |

| | |
|------------------------|---|
| Dariusz Mrozek | Silesian University of Technology, Poland |
| Richi Nayak | Queensland University of Technology, Australia |
| Paulo Novais | Universidade do Minho, Portugal |
| Boris Novikov | Saint Petersburg University, Russian Federation |
| Jisha Jose Panackal | Vidya Academy of Science and Technology, India |
| George Papastefanatos | Athena Research and Innovation Center, Greece |
| Jeffrey Parsons | Memorial University of Newfoundland, Canada |
| Barbara Pernici | Politecnico di Milano, Italy |
| Ilia Petrov | Reutlingen University, Germany |
| Iulian Sandu Popa | University of Versailles Saint-Quentin-en-Yvelines and Inria Saclay, France |
| Nirvana Popescu | University Politehnica of Bucharest, Romania |
| Philippe Pucheral | University of Versailles Saint-Quentin en Yvelines (UVSQ), France |
| Elisa Quintarelli | Politecnico di Milano, Italy |
| Christoph Quix | RWTH Aachen University, Germany |
| Praveen Rao | University of Missouri-Kansas City, USA |
| Alexander Rasin | DePaul University, USA |
| Kun Ren | Yale University, USA |
| Colette Rolland | Université De Paris I Panthéon Sorbonne, France |
| Gustavo Rossi | Lifia, Argentina |
| Gunter Saake | Institute of Technical and Business Information Systems, Germany |
| Dimitris Sacharidis | Technische Universität Wien, Austria |
| Manuel Filipe Santos | Centro ALGORITMI, University of Minho, Portugal |
| Maria Luisa Sapino | Università di Torino, Italy |
| Ralf Schenkel | University of Trier, Germany |
| Diego Seco | University of Concepción, Chile |
| Vinay Setty | Max Planck Institut für Informatik, Germany |
| Ijün shan | Inria, France |
| Nematollaah Shiri | Concordia University, Canada |
| Harvey Siy | University of Nebraska at Omaha, USA |
| Spiros Skiadopoulos | University of Peloponnese, Greece |
| Yeong-Tae Song | Towson University, USA |
| Sergey Stupnikov | IPI RAN, Russian Federation |
| Zbigniew Suraj | University of Rzeszow, Poland |
| Nemat El Tazi | Cairo University, Egypt |
| Maguelonne Teisseire | Irstea, National Research Institute of Science and Technology for Environment and Agriculture, France |
| Manolis Terrovitis | Institute for the Management of Information Systems, Greece |
| Babis Theodoulidis | University of Manchester, UK |
| Frank Tompa | University of Waterloo, Canada |
| Christos Tryfonopoulos | University of Peloponnese, Greece |

| | |
|--------------------|---|
| Goran Velinov | UKIM, Macedonia, Former Yugoslav Republic of |
| Thanasis Vergoulis | Athena Research and Innovation Center, Greece |
| Karin Verspoor | University of Melbourne, Australia |
| José Ríos Viqueira | Universidade de Santiago de Compostela, Spain |
| Gianluigi Viscusi | EPFL Lausanne, Switzerland |
| Hannes Voigt | TU Dresden, Germany |
| Florian Wenzel | University of Augsburg, Germany |
| Leandro Krug Wives | Universidade Federal do Rio Grande do Sul, Brazil |
| Robert Wrembel | Poznan University of Technology, Poland |
| Yun Xiong | Fudan University, China |
| Filip Zavoral | Charles University Prague, Czech Republic |
| Jiakui Zhao | State Grid Information and Telecommunication Group of China, China |
| Jianlong Zhong | GRAPHSQL Inc., USA |
| Yangyong Zhu | Fudan University, China |

Additional Reviewers

| | |
|-----------------------|--|
| Idir Benouaret | Heudiasyc Laboratory, UMR CNRS 7253, Université de Technologie de Compiègne, France |
| Dominik Bork | University of Vienna, Austria |
| Michele A. Brandao | Universidade Federal de Minas Gerais, Brazil |
| Estrela Ferreira Cruz | Instituto Politécnico de Viana do Castelo, Portugal |
| Giacomo Domeniconi | University of Bologna, Italy |
| José Fuentes | University of Concepción, Chile |
| Olga Gkountouna | Research Center Athena, Greece |
| Emanuele Rabosio | Politecnico di Milano, Italy |

Invited Speakers

| | |
|-------------------|---------------------------------|
| Panos Vassiliadis | University of Ioannina, Greece |
| Christoph Quix | RWTH Aachen University, Germany |

Contents

| | |
|---|-----|
| A Scalable Platform for Low-Latency Real-Time Analytics of Streaming Data | 1 |
| <i>Paolo Cappellari, Mark Roantree, and Soon Ae Chun</i> | |
| Identifying Conversational Message Threads by Integrating Classification and Data Clustering. | 25 |
| <i>Giacomo Domeniconi, Konstantinos Semertzidis, Gianluca Moro, Vanessa Lopez, Spyros Kotoulas, and Elizabeth M. Daly</i> | |
| Towards Efficient Multi-domain Data Processing | 47 |
| <i>Johannes Luong, Dirk Habich, Thomas Kissinger, and Wolfgang Lehner</i> | |
| Approaching ETL Processes Specification Using a Pattern-Based Ontology . . . | 65 |
| <i>Bruno Oliveira and Orlando Belo</i> | |
| Topic-Aware Visual Citation Tracing via Enhanced Term Weighting for Efficient Literature Retrieval | 79 |
| <i>Youbing Zhao, Hui Wei, Shaopeng Wu, Farzad Parvinzamid, Zhikun Deng, Xia Zhao, Nikolaos Ersotelos, Feng Dong, Gordon Clapworthy, and Enjie Liu</i> | |
| Maturing Pay-as-you-go Data Quality Management: Towards Decision Support for Paying the Larger Bills. | 102 |
| <i>Jan van Dijk, Mortaza S. Bargh, Sunil Choenni, and Marco Spruit</i> | |
| Generic and Concurrent Computation of Belief Combination Rules. | 125 |
| <i>Frédéric Dambreville</i> | |
| Log-Based Model to Enforce Data Consistency on Agnostic Fault-Tolerant Systems | 141 |
| <i>Óscar Mortágua Pereira, David Apolinário Simões, and Rui L. Aguiar</i> | |
| Improving Performances of an Embedded Relational Database Management System with a Hybrid CPU/GPU Processing Engine. | 160 |
| <i>Samuel Cremer, Michel Bagein, Saïd Mahmoudi, and Pierre Manneback</i> | |
| Author Index | 179 |