

Slimane Hammoudi
Christoph Quix
Jorge Bernardino (Eds.)

Communications in Computer and Information Science

1446

Data Management Technologies and Applications

9th International Conference, DATA 2020
Virtual Event, July 7–9, 2020
Revised Selected Papers

Editorial Board Members

Joaquim Filipe 

Polytechnic Institute of Setúbal, Setúbal, Portugal

Ashish Ghosh

Indian Statistical Institute, Kolkata, India

Raquel Oliveira Prates 

Federal University of Minas Gerais (UFMG), Belo Horizonte, Brazil

Lizhu Zhou

Tsinghua University, Beijing, China

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

Slimane Hammoudi · Christoph Quix ·
Jorge Bernardino (Eds.)

Data Management Technologies and Applications

9th International Conference, DATA 2020
Virtual Event, July 7–9, 2020
Revised Selected Papers

Editors

Slimane Hammoudi
MODESTE/ESEO
Angers, France

Christoph Quix
Fraunhofer FIT and RWTH Aachen
University
Aachen, Germany

Jorge Bernardino
University of Coimbra
Coimbra, Portugal

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-030-83013-7 ISBN 978-3-030-83014-4 (eBook)
<https://doi.org/10.1007/978-3-030-83014-4>

© Springer Nature Switzerland AG 2021

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, expressed 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.

This Springer imprint is published by the registered company Springer Nature Switzerland 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 9th International Conference on Data Science, Technology and Applications (DATA 2020), exceptionally held as a web-based event, due to the COVID-19 pandemic, during July 7–9, 2020.

DATA 2020 received 70 paper submissions from 31 countries, of which 20% were 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 purpose of the 9th International Conference on Data Science, Technology and Applications (DATA) was to bring together researchers, engineers and practitioners interested in databases, big data, 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 data science, technology and applications, especially of approaches required to tackle current and future challenges in data science and data management. Thus, this book covers diverse but complementary topics such as data management and quality, new computational models for big data, big data search and mining, statistics exploratory data analysis, predictive modelling, big data infrastructure and architecture, data privacy and security, data and big data in Industry 4.0.

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

July 2020

Slimane Hammoudi
Christoph Quix
Jorge Bernardino

Organization

Conference Chair

Jorge Bernardino Polytechnic of Coimbra - ISEC, Portugal

Program Co-chairs

Slimane Hammoudi ESEO, France
Christoph Quix Hochschule Niederrhein, University of Applied Sciences/Fraunhofer FIT, Germany

Program Committee

Maha Amami University of Milan-Bicocca, Italy
Christos Anagnostopoulos University of Glasgow, UK
Gustavo Arroyo-Figueroa Instituto Nacional de Electricidad y Energías Limpias, Mexico
Karim Benouaret Université Claude Bernard Lyon 1, France
Jan Bohacik University of Zilina, Slovak Republic
Gloria Bordogna ICAR-CNR, Italy
Nieves R. Brisaboa University of A Coruña, Spain
Cinzia Cappiello Politecnico di Milano, Italy
Paola Carrara IREA-CNR, Italy
Richard Chbeir Université de Pau et des Pays de l'Adour, France
Antonio Corral University of Almeria, Spain
Gianni Costa ICAR-CNR, Italy
Shruti Daggumati University of Nebraska-Lincoln, USA
Theodore Dalamagas Athena Research and Innovation Center, Greece
Bruno Defude Institut Mines Telecom, France
Steven Demurjian University of Connecticut, USA
Martin Drlik Constantine the Philosopher University in Nitra, Slovak Republic
Fabien Duchateau Université Claude Bernard Lyon 1/LIRIS, France
John Easton University of Birmingham, UK
Markus Endres University of Augsburg, Germany
Francesco Folino ICAR-CNR, Italy
Pedro Furtado University of Coimbra, Portugal
Jérôme Gensel Université Grenoble Alpes, France
Paola Giannini University of Piemonte Orientale, Italy
John Gibson Telecom SudParis, France
Janis Grabis Riga Technical University, Latvia

Francesco Guerra	University of Modena and Reggio Emilia, Italy
Rihan Hai	RWTH Aachen University, Germany
Raju Halder	Indian Institute of Technology Patna, India
Andreas Henrich	University of Bamberg, Germany
Jose Herrera	Universidad del Cauca, Colombia
Jang-Eui Hong	Chungbuk National University, South Korea
Tsan-Sheng Hsu	Institute of Information Science, Academia Sinica, Taiwan, Republic of China
Sergio Ilarri	University of Zaragoza, Spain
Ivan Ivanov	SUNY Empire State College, USA
Nikos Karacapilidis	University of Patras, Greece
Pawel Kasprzycki	Silesian University of Technology, Poland
Takahiro Kawamura	Japan Science and Technology Agency, Japan
David Kenschke	SAP, Germany
Angeliki Kitsiou	University of the Aegean, Greece
Kostas Kolomvatos	National and Kapodistrian University of Athens, Greece
Martin Krulis	Charles University in Prague, Czech Republic
Vladimir Kurbalija	University of Novi Sad, Serbia
Jean-Charles Lamirel	Loria/University of Strasbourg, France
Wolfgang Lehner	Technische Universität Dresden, Germany
Raimondas Lencevicius	Nuance Communications, USA
Christos Makris	University of Patras, Greece
Yannis Manolopoulos	Open University of Cyprus, Cyprus
Miguel Martínez-Prieto	University of Valladolid, Spain
Florent Masseglia	Inria, France
Amin Mesmoudi	Université de Poitiers, France
Yasser Mohammad	Assiut University, Egypt
Stefano Montanelli	Università degli Studi di Milano, Italy
Mirella Moro	Federal University of Minas Gerais, Brazil
Erich Neuhold	University of Vienna, Austria
Riccardo Ortale	ICAR-CNR, Italy
George Papastefanatos	Athena Research and Innovation Center, Greece
Jeffrey Parsons	Memorial University of Newfoundland, Canada
Luca Piovesan	Università del Piemonte Orientale, Italy
Nirvana Popescu	University Politehnica of Bucharest, Romania
Dimitrios Rafailidis	Maastricht University, The Netherlands
Paraskevi Raftopoulou	University of the Peloponnese, Greece
Werner Retschitzegger	Johannes Kepler University, Austria
José Ríos Viqueira	Universidade de Santiago de Compostela, Spain
Colette Rolland	Université Paris 1 Panthéon Sorbonne, France
Gustavo Rossi	Lifia, Argentina
Mercedes Ruiz	University of Cadiz, Spain
Gunter Saake	Institute of Technical and Business Information Systems, Germany
Dimitris Sacharidis	Technische Universität Wien, Austria

Iulian Sandu Popa	University of Versailles Saint-Quentin-en-Yvelines/Inria, France
Diego Seco	University of Concepción, Chile
Nematollah Shiri	Concordia University, Canada
Marius Silaghi	Florida Institute of Technology, USA
Stavros Simou	University of the Aegean, Greece
Spiros Skiadopoulos	University of the Peloponnese, Greece
Cosmin Stoica	Romania
Sergey Stupnikov	IPI RAN, Russia
Zbigniew Suraj	University of Rzeszow, Poland
George Tambouratzis	Institute for Language and Speech Processing, Greece
Tatiana Tambouratzis	University of Piraeus, Greece
Horia-Nicolai Teodorescu	Gheorghe Asachi Technical University of Iasi, Romania
Catarci Tiziana	Università degli Studi di Roma La Sapienza, Italy
Maurice van Keulen	University of Twente, The Netherlands
Michael Vassilakopoulos	University of Thessaly, Greece
Thanasis Vergoulis	Athena Research and Innovation Center, Greece
Marco Villani	University of Modena and Reggio Emilia, Italy
Gianluigi Viscusi	EPFL Lausanne, Switzerland
Zeev Volkovich	Ort Braude College, Israel
Leandro Wives	Universidade Federal do Rio Grande do Sul, Brazil
Shengkun Xie	Ryerson University, Canada
Filip Zavoral	Charles University in Prague, Czech Republic
Jiakui Zhao	State Grid Big Data Center of China, China

Additional Reviewers

Michele A. Brandão	IFMG, Brazil
Karam Bou Chaaya	Spider Research Group, France
Athanasios Davvetas	University of the Peloponnese, Greece
Abbas Javadtalab	Concordia University, Canada
Susana Ladra Gonzalez	University of A Coruña, Spain
Stavros Maroulis	Athena Research and Innovation Center, Greece

Invited Speakers

Wil van der Aalst	RWTH Aachen University, Germany
Ioana Manolescu	Inria, France

Contents

Removing Operational Friction Using Process Mining: Challenges Provided by the Internet of Production (IoP)	1
<i>Wil M. P. van der Aalst, Tobias Brockhoff, Anahita Farhang Ghahfarokhi, Mahsa Pourbafrani, Merih Seran Uysal, and Sebastiaan J. van Zelst</i>	
Efficient Scheduling of Scientific Workflow Actions in the Cloud Based on Required Capabilities	32
<i>Michel Krämer</i>	
iTLM-Q: A Constraint-Based Q-Learning Approach for Intelligent Traffic Light Management	56
<i>Christian Roth, Lukas Stöger, Mirja Nitschke, Matthias Hörmann, and Dogan Kesdogan</i>	
Open Data in the Enterprise Context: Assessing Open Corporate Data's Readiness for Use	80
<i>Pavel Krasikov, Timo Obrecht, Christine Legner, and Markus Eurich</i>	
A Data Science Approach to Explain a Complex Team Ball Game	101
<i>Friedemann Schwenkreis</i>	
Intelligent Public Procurement Monitoring System Powered by Text Mining and Balanced Indicators	115
<i>Nikola Modrušan, Leo Mršić, and Kornelije Rabuzin</i>	
Catalog Integration of Heterogeneous and Volatile Product Data	134
<i>Oliver Schmidts, Bodo Kraft, Marvin Winkens, and Albert Zündorf</i>	
Designing an Efficient Gradient Descent Based Heuristic for Clusterwise Linear Regression for Large Datasets	154
<i>Enis Kayış</i>	
A Policy-Agnostic Programming Language for the International Data Spaces	172
<i>Fabian Bruckner, Julia Pampus, and Falk Howar</i>	
Coreset-Based Data Compression for Logistic Regression	195
<i>Nery Riquelme-Granada, Khuong An Nguyen, and Zhiyuan Luo</i>	

Product Classification Using Partially Abbreviated Product Names, Brands and Dimensions	223
<i>Oliver Allweyer, Christian Schorr, Andreas Mohr, and Rolf Krieger</i>	
Algebraic Expressions with State Constraints for Causal Relations and Data Semantics	245
<i>Susumu Yamasaki and Mariko Sasakura</i>	
An Environmental Study of French Neighbourhoods	267
<i>Nelly Barret, Fabien Duchateau, Franck Favetta, Aurélien Gentil, and Loïc Bonneval</i>	
Phenomena Explanation from Text: Unsupervised Learning of Interpretable and Statistically Significant Knowledge	293
<i>Giacomo Frisoni and Gianluca Moro</i>	
Author Index	319