

## Data-Driven Process Discovery and Analysis

Paolo Ceravolo, Maurice van Keulen, María Teresa Gómez-López

► **To cite this version:**

Paolo Ceravolo, Maurice van Keulen, María Teresa Gómez-López. Data-Driven Process Discovery and Analysis. Springer International Publishing, LNBIP-379, 2020, Lecture Notes in Business Information Processing, 978-3-030-46632-9. 10.1007/978-3-030-46633-6 . hal-03188665

**HAL Id: hal-03188665**

**<https://hal.inria.fr/hal-03188665>**

Submitted on 2 Apr 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Lecture Notes in Business Information Processing

379

## Series Editors

Wil van der Aalst 

*RWTH Aachen University, Aachen, Germany*

John Mylopoulos 

*University of Trento, Trento, Italy*

Michael Rosemann 

*Queensland University of Technology, Brisbane, QLD, Australia*

Michael J. Shaw

*University of Illinois, Urbana-Champaign, IL, USA*

Clemens Szyperski

*Microsoft Research, Redmond, WA, USA*

More information about this series at <http://www.springer.com/series/7911>


Paolo Ceravolo · Maurice van Keulen ·  
María Teresa Gómez-López (Eds.)


# Data-Driven Process Discovery and Analysis

8th IFIP WG 2.6 International Symposium, SIMPDA 2018  
Seville, Spain, December 13–14, 2018  
and 9th International Symposium, SIMPDA 2019  
Bled, Slovenia, September 8, 2019  
Revised Selected Papers

*Editors*

Paolo Ceravolo   
Università degli Studi di Milano  
Milan, Italy

Maurice van Keulen   
University of Twente  
Enschede, The Netherlands

María Teresa Gómez-López   
University of Seville  
Seville, Spain

ISSN 1865-1348                      ISSN 1865-1356 (electronic)  
Lecture Notes in Business Information Processing  
ISBN 978-3-030-46632-9              ISBN 978-3-030-46633-6 (eBook)  
<https://doi.org/10.1007/978-3-030-46633-6>

© IFIP International Federation for Information Processing 2020

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 rapid growth at which organizations and businesses process data, managed via information systems, has made available a big variety of information that consequently has created a high demand for making data analytics more effective and valuable. Today, these new data analyzing techniques have to cope with the continuous advancements in the digital transformation course. Blockchain infrastructures bring trusted transactions to interorganizational procedures. The growing maturity level of Artificial Intelligence solutions conveys the integration of various analyzing techniques and cultures. IoT technologies bring traceability to potentially any human-in-the-loop process. The eighth and ninth editions of the International Symposium on Data-driven Process Discovery and Analysis (SIMPDA 2018, 2019) were conceived to offer a forum where researchers from different communities could share their insights in this hot new field. As a symposium, SIMPDA fosters exchanges among academic researchers, industry experts, and a wider audience interested in process discovery and analysis.

Submissions cover theoretical issues related to processing representation, discovery, and analysis, or provide practical and operational examples of their application. To improve the quality of the contributions, the symposium is structured towards fostering discussion and stimulating improvements. Papers are pre-circulated to the authors, who are expected to read them and make ready comments and suggestions. After the event, authors have the opportunity to improve their work by extending the presented results. For this reason, authors of accepted papers were invited to submit extended articles to this post-symposium volume. We received 25 submissions and 6 papers were accepted for publication.

The current selection of papers underlines the most relevant challenges that were identified, and proposes novel solutions for facing these challenges.

In the first paper, “Designing Process-Centric Blockchain-based Architectures: A Case Study in e-voting as a Service,” Emanuele Bellini et al. study a solution to put into operation Business Process Management on a Blockchain-based infrastructure to develop a diverse set of execution and monitoring systems on Blockchain, and define appropriate methods for evolution and adaptation.

The second paper, by Berti and Van der Aalst, is titled “Extracting Multiple Viewpoint Models from Relational Databases,” and presents an advanced methodology for coping with discovering multiple viewpoints in relational databases collecting event log data.

The third paper by Cancela et al., “Standardizing Process-Data Exploitation by means of a Process-Instance Metamodel,” proposes the use of a Business-Process Instance Metamodel as an interface between the applications that produce the data and those which consume the data making their data structures independent and, by consequence, their set up less expensive.

The fourth paper by Hinkka et al., “Exploiting Event Log Event Attributes in RNN Based on Prediction,” discusses a method for giving a trade-off between prediction accuracy and training time in RNN predictions over business process cases. This trade-off is achieved by enriching the vectors imputing the RNN using the groups obtained by clustering techniques. This experimental analysis shows that having event attribute clusters encoded into the input vectors outperforms having the actual attribute values in the input vector.

The fifth paper by Martinez-Gil et al., “General Model for Tracking Manufacturing Products Using Graph Databases,” presents a case study on product manufacturing where a graph database is exploited to reduce response times in tracking the process execution.

The sixth paper by Rafei et al., “Supporting Confidentiality in Process Mining Using Abstraction and Encryption,” presents an approach for supporting data encryption in Process Mining. Using abstraction, the authors hide confidential information in a controlled manner while ensuring that the desired Process Mining results can still be obtained.

We gratefully acknowledge the research community that gathered around the problems related to process data analysis. We would also like to express our deep appreciation for the referees’ hard work and dedication. Above all, thanks are due to the authors for submitting the best results of their work to SIMPDA.

We are very grateful to the Università degli Studi di Milano and IFIP for their organizational support.

March 2020

Paolo Ceravolo  
Maurice van Keulen  
María Teresa Gómez-López

# Organization

## Chairs

Paolo Ceravolo	Università degli Studi di Milano, Italy
Maurice Van Keulen	University of Twente, The Netherlands
Maria Teresa Gomez Lopez	University of Seville, Spain

## Advisory Board

Ernesto Damiani	Università degli Studi di Milano, Italy
Erich Neuhold	University of Vienna, Austria
Philippe Cudré-Mauroux	University of Fribourg, Switzerland
Robert Meersman	Graz University of Technology, Austria
Wilfried Grossmann	University of Vienna, Austria

## SIMPDA Award Committee

Maria Teresa Gomez Lopez	University of Seville, Spain
Paolo Ceravolo	Università degli Studi di Milano, Italy

## Web and Publicity Chair

Fulvio Frati	Università degli Studi Di Milano, Italy
--------------	---

## Program Committee

Alexandra Mazak	University of Vienna, Austria
Robert Singer	FH Joanneum, Austria
Manfred Reichert	University of Ulm, Germany
Schahram Dustdar	Vienna University of Technology, Austria
Helen Balinsky	Hewlett-Packard Laboratories, UK
Valentina Emilia Balas	University of Arad, Romania
Antonio Mana Gomez	Universidad de Málaga, Spain
Karima Boudaoud	École polytechnique de l'université Nice-Sophia Antipolis, France
Jan Mendling	Vienna University of Economics and Business, Austria
Peter Spyns	Flemish Government, Belgium
Mohamed Mosbah	University of Bordeaux, France
Chintan Mrit	University of Twente, The Netherlands
Fabrizio Maria Maggi	University of Tartu, Estonia
Pnina Soffer	University of Haifa, Israel
Matthias Weidlich	Imperial College, UK



Roland Rieke	Fraunhofer Sit, Germany
Edgar Weippl	Vienna University of Technology, Austria
Benoit Depaire	University of Hasselt, Belgium
Angel Jesus	Varela University of Seville, Spain
Luisa Parody	University Loyola Andalucia, Spain
Antonia Azzini	Consorzio C2t, Italy
Jorge Cardoso	University of Coimbra, Spain
Carlos Fernandez-Llatas	Universitat Politècnica de València, Spain
Chiara di Francescomarino	Fondazione Bruno Kessler, Italy
Faiza Bukhsh	University of Twente, The Netherlands
Mirjana Pejić Bach	University of Zagreb, Croatia
Tamara Quaranta	40labs, Italy
Anna Wilbik	Eindhoven University of Technology, The Netherlands
Yingqian Zhang	Eindhoven University of Technology, The Netherlands
Richard Chbeir	Université de Pau et des Pays de l'Adour, France
Renata Medeiros	Eindhoven University of Technology, The Netherlands
Rabia Maqsood	Università degli Studi di Milano, Italy

# Contents

Designing Process-Centric Blockchain-Based Architectures: A Case Study in e-voting as a Service . . . . .	1
<i>Emanuele Bellini, Paolo Ceravolo, Alessandro Bellini, and Ernesto Damiani</i>	
Extracting Multiple Viewpoint Models from Relational Databases. . . . .	24
<i>Alessandro Berti and Wil van der Aalst</i>	
Standardizing Process-Data Exploitation by Means of a Process-Instance Metamodel . . . . .	52
<i>Antonio Cancela, Antonia M. Reina Quintero, María Teresa Gómez-López, and Alejandro García-García</i>	
Exploiting Event Log Event Attributes in RNN Based Prediction . . . . .	67
<i>Markku Hinkka, Teemu Lehto, and Keijo Heljanko</i>	
General Model for Tracking Manufacturing Products Using Graph Databases . . . . .	86
<i>Jorge Martinez-Gil, Reinhard Stumptner, Christian Lettner, Mario Pichler, Salma Mahmoud, Patrick Praher, and Bernhard Freudenthaler</i>	
Supporting Confidentiality in Process Mining Using Abstraction and Encryption . . . . .	101
<i>Majid Rafiei, Leopold von Waldthausen, and Wil M. P. van der Aalst</i>	
<b>Author Index</b> . . . . .	125