



# **DATA ANALYTICS 2016**

The Fifth International Conference on Data Analytics

October 9 - 13, 2016

Venice, Italy

## **DATA ANALYTICS 2016 Editors**

Sandjai Bhulai, Vrije Universiteit Amsterdam, the Netherlands

Ivana Semanjski, University of Zagreb, Croatia / Ghent University, Belgium

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2016) by International Academy, Research, and Industry Association (IARIA)  
Please refer to the Copyright Information page.

Printed by Curran Associates, Inc. (2016)

International Academy, Research, and Industry Association (IARIA)  
412 Derby Way  
Wilmington, DE 19810

Phone: (408) 893-6407  
Fax: (408) 527-6351

[petre@iaria.org](mailto:petre@iaria.org)

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# TABLE OF CONTENTS

## DATA ANALYTICS 1: TARGET ANALYTICS

<b>Performance of Spanish Encoding Functions During Record Linkage</b> .....	1
<i>María Del Pilar Angeles, Noemi Bailón-Miguel</i>	
<b>Leveraging Analytics to Predict Geomagnetic Storms: Impact to Global Telecommunications</b> .....	8
<i>Taylor K. Larkin, Denise J. McManus</i>	
<b>A Framework for Call Center Decongestion Using Sequential Pattern Analysis</b> .....	14
<i>Eugene Rex L. Jalao</i>	
<b>Tracking Cyclists and Walkers: Will it Change Planning and Policy Processes?</b> .....	18
<i>João Bernardino, Mafalda Lopes, Predrag Živanovic, Slaven Tica, Branko Milovanovic, Stanko Bajcetic</i>	

## DATA ANALYTICS 2: FUNDAMENTALS

<b>The 100-fold Cross Validation for Small Sample Method</b> .....	29
<i>Shuichi Shinmura</i>	
<b>Improving Process Mining Prediction Results in Processes that Change Over Time</b> .....	37
<i>Alessandro Berti</i>	
<b>Real-Time Knowledge Map Services on National R&amp;D Data</b> .....	43
<i>Kang-Ryul Shon, Cheol-Joo Chae, Han-Jo Jeong, Chul-Su Lim</i>	

## DATA ANALYTICS 3: SENTIMENT/OPINION ANALYSIS

<b>Multilingual Sentiment Analysis on Data of the Refugee Crisis in Europe</b> .....	45
<i>Gayane Shalunts, Gerhard Backfried</i>	
<b>The Impact of Machine Translation on Sentiment Analysis</b> .....	51
<i>Gayane Shalunts, Gerhard Backfried, Nicolas Commeignes</i>	
<b>Car Sales Forecasting Using Artificial Neural Networks and Analytical Hierarchy Process – Case Study: Kia and Hyundai Corporations in the USA</b> .....	57
<i>Danial Shahrabi Farahani, Mansour Momeni, Nader Sayyed Amiri</i>	

## DATA ANALYTICS 4: APPLICATION-ORIENTED ANALYSIS

<b>Predicting Candidate Uptake For Online Vacancies</b> .....	63
<i>Corne De Ruijt, Sandjai Bhulai, Han Rusman, Leon Willemsens</i>	
<b>Almost Squares in Almost Squares: Solving the Final Instance</b> .....	69
<i>Daan Van Den Berg, Florian Braam, Mark Moes, Emiel Suilen, Sandjai Bhulai</i>	
<b>Twitter Analytics for the Horticulture Industry</b> .....	75
<i>Marijn Ten Thij, Sandjai Bhulai, Wilco Van Den Berg, Henk Zwinkels</i>	
<b>Using Data Mining Techniques for Information System Research Purposes – An Exemplary Application in the Field of Business Intelligence and Corporate Performance Management Research</b> .....	80
<i>Karin Hartl, Olaf Jacob</i>	

## DATA ANALYTICS 5 / TTASC: TRANSPORT AND TRAFFIC ANALYTICS IN SMART CITIES

<b>Travelled Distance Estimation for GPS-based Round Trips: Car-Sharing Use Case</b> .....	87
<i>Angel J. Lopez, Ivana Semanjski, Dominique Gillis, Daniel Ochoa, Sidharta Gautama</i>	
<b>Forecasting Travel Behaviour from Crowdsourced Data with Machine Learning Based Model</b> .....	93
<i>Angel J. Lopez, Ivana Semanjski, Sidharta Gautama</i>	
<b>Datamining and Big Freight Transport Database: Analysis and Forecasting Capabilities</b> .....	100
<i>Massimiliano Petri, Antonio Pratelli, Giovanni Fusco</i>	

<b>Seaplane Traffic in the Republic of Croatia</b> .....	109
<i>Pero Vidan, Merica Sliškovic, Nikola Ocašić</i>	

## **DATA ANALYTICS 6: BIG DATA**

<b>WACIC Method – A Web Analytics Process to Perform Continuous Improvement in Digital Environments</b> .....	118
<i>Adriana T. Figueiredo, Marcos Augusto F. Borges, Regina Lucia De O. Moraes</i>	
<b>Learning Analytics: Supporting Teaching and Learning through Learner’s Data Analytics and Visualization</b> .....	126
<i>Ali Shiri</i>	
<b>Tourism Websites Network: Crawling the Italian Webspaces</b> .....	131
<i>Alessandro Longheu, Giuseppe Mangioni, Marialaura Previti</i>	
<b>A Study of the OAuth 2.0 Protocol Extended Using SMS for Safe User Access</b> .....	137
<i>Chae Cheol-Joo, Kwang-Nam Choi</i>	
<b>Author Index</b>	