

Communications in Computer and Information Science

898

Commenced Publication in 2007

Founding and Former Series Editors:

Phoebe Chen, Alfredo Cuzzocrea, Xiaoyong Du, Orhun Kara, Ting Liu,
Dominik Ślęzak, and Xiaokang Yang

Editorial Board

Simone Diniz Junqueira Barbosa

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

Joaquim Filipe

Polytechnic Institute of Setúbal, Setúbal, Portugal

Ashish Ghosh

Indian Statistical Institute, Kolkata, India

Igor Kotenko

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

Krishna M. Sivalingam

Indian Institute of Technology Madras, Chennai, India

Takashi Washio

Osaka University, Osaka, Japan

Junsong Yuan

University at Buffalo, The State University of New York, Buffalo, USA

Lizhu Zhou

Tsinghua University, Beijing, China


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


Juan Antonio Lossio-Ventura
Denisse Muñante · Hugo Alatrasta-Salas (Eds.)


Information Management and Big Data

5th International Conference, SIMBig 2018
Lima, Peru, September 3–5, 2018
Proceedings

Editors

Juan Antonio Lossio-Ventura 
Biomedical Informatics,
College of Medicine
University of Florida
Gainesville, FL, USA

Hugo Alatrística-Salas 
Facultad de Ingeniería
University of the Pacific
Jesús María, Lima, Peru

Denisse Muñante 
Fondazione Bruno Kessler
Trento, Italy

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-3-030-11679-8 ISBN 978-3-030-11680-4 (eBook)
<https://doi.org/10.1007/978-3-030-11680-4>

Library of Congress Control Number: 2018968324

© Springer Nature Switzerland AG 2019

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.

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

Preface

Today, data scientists use the term “big data” to describe the exponential growth and availability of data, which could be structured and unstructured. Big data has taken place over the past 20 years and will stay with us for the next few years. In a recent study, Cisco Systems predicted that in 2021, over 27.1 billion electronic devices would be connected to the Internet. From these devices, smartphones only will generate — on average — 14.9 gigabytes of data per month. This massive amount of data contains, possibly, patterns of behavior that can be exploited by organizations to take decisions or to design public policies with a high impact. In this context, the term big data not only concerns the storage, the management, and the analysis of a large amount of data but also is associated with the ability to implement new algorithms, to propose new techniques, to apply new strategies, and to find new real-world applications.

Several domains, including biomedicine, life sciences, and scientific research, have been affected by big data. For instance, social networks such as Facebook, Twitter, and LinkedIn generate masses of data, which are available to be accessed by other applications. Therefore, there is a need to understand and exploit all kinds of data (structured and unstructured). This process can be performed with data science, which encompasses methods of data mining, natural language processing, the Semantic Web, and statistics, among others, which allow us to gain new insights from data.

SIMBig (International Conference on Information Management and Big Data) seeks to present new methods of the fields related to data science to assess large volumes of data. SIMBig aims to bring together main — national and international — actors in the decision-making field to state in new technologies dedicated to handling amounts of data. Besides, the conference is a convivial place where these actors can present their innovative contributions and receive feedback from the experts.

This book collects three invited talks and 30 accepted contributions from 101 submitted papers belonging to the following four special tracks:

- ANLP - Track on Applied Natural Language Processing
- DISE - Track on Data-Driven Software Engineering
- SNMAM - Track on Social Network and Media Analysis and Mining
- PSCBig - Track on Privacy and Security Challenges on Big Data

SIMBig 2018 had six keynote speakers who are experts in the main topics of the conference. The conference started with the talk given by our invited speaker Andrei Broder. Broder, a distinguished scientist at Google, reviewed the evolution of Internet search engines, highlighting that his success and survival is centered on three axes: scalability, speed, and functionality. Broder said that the last big step of Google is the jump from the search drawer to the personal assistant. He also stressed that the technology of Internet search is based on the user experience and that the life cycle of innovations is very short, since “the new” quickly becomes “normal.” “If you looked at the technology of the search when the autocomplete was introduced in the query box,

this was a breakthrough; people were surprised to be able to type only a couple of letters and get what they wanted. Now, if you have a query box and do not auto-complete, the user is annoyed. Surely very soon if you have to write will generate a total annoyance, because users will want to have voice recognition,” said Broder exemplifying the virtual circle of innovation.

Furthermore, Lise Getoor from the University of California presented a talk entitled “Effectively Exploiting Structure in Data Science Problems,” which is summarized as: “Our ability to collect, manipulate, analyze, and act on vast amounts of data is having a profound impact on all aspects of society. Much of this data is heterogeneous in nature and interlinked in a myriad of complex ways. From information integration to scientific discovery to computational social science, we need machine learning methods that are able to exploit both the inherent uncertainty and the innate structure in a domain. Statistical relational learning (SRL) is a machine learning subfield that builds on principles from probability theory and statistics to address uncertainty while incorporating tools from knowledge representation and logic to represent structure.” Getoor overviewed her recent work on probabilistic soft logic (PSL), an SRL framework for large-scale collective, probabilistic reasoning in relational domains. PSL provides a tractable approach to probabilistic inference. She showed the theoretical foundations for PSL, which connect work from the theoretical computer science community on randomized algorithms with work from the probabilistic graphical model’s community on local consistency relaxations with work from the AI community on soft logic. She also described several successful applications of PSL to problems from computational social science (stance in online forums, social trust, latent political groups, cyberbullying) and data integration (entity resolution and knowledge graph construction). Getoor closed up with a discussion of responsible data science, which requires understanding both the inherent structure in the domain, the structure and potential bias in the data, and the potential implications and feedback loops in the algorithms.

Later, Jian Pei, vice-president of JD.com and professor at the Simon Fraser University, gave a talk about data mining and logistic. The talk is summarized as: “The future of retail is in breaking the limitations of the current business models in customer connection, logistics and stores. The key to breaking those limitations is the integrated smart supply chain, which covers smart consumption, smart logistics, and smart supply. The foundation of a smart supply chain is intelligent big data science. Through a series of examples, we discuss how smart supply chain techniques and data science can enable the future of retail. Specifically, we demonstrate how big data and AI techniques together can deepen our understanding of customers, create new convenience and efficiency in retail scenarios, minimize the cost store operation, and shorten the path from customers to manufacturers”.

Finally, the remaining three invited talks were presented in the form of short papers in this book.

To share the new analysis methods for managing large volumes of data, we encouraged participation from researchers in all fields related to big data, data science, data mining, natural language processing, and Semantic Web, but also multilingual text processing, biomedical NLP, data-driven software engineering, and data privacy.

Topics of interest to SIMBig included: data science, big data, data mining, natural language processing, bio NLP, text mining, information retrieval, machine learning,

Semantic Web, ontologies, IoT, privacy on social networks, Web mining, knowledge representation and linked open data, social networks, social Web, and Web science, information visualization, OLAP, data warehousing, business intelligence, spatiotemporal data, health care, agent-based systems, data-driven security, and privacy.

SIMBig is positioning itself as one of the most important conferences in South America on issues related to information management and Big Data. Two Springer CCIS books were published in the context of the SIMBig conference. The first publication compiles the selected papers from the SIMBig 2015 and SIMBig 2016 conferences [1]. The second one presents the best papers of the SIMBig 2017 conference [2].

References

1. J. A. Lossio-Ventura and H. Alatrística-Salas, editors. *Information Management and Big Data - Second Annual International Symposium, SIMBig 2015, Cusco, Peru, September 2–4, 2015, and Third Annual International Symposium, SIMBig 2016, Cusco, Peru, September 1–3, 2016, Revised Selected Papers*, volume 656 of *Communications in Computer and Information Science*. Springer, 2017.
2. J. A. Lossio-Ventura and H. Alatrística-Salas, editors. *Information Management and Big Data - 4th Annual International Symposium, SIMBig 2017, Lima, Peru, September 4–6, 2017, Revised Selected Papers*, volume 795 of *Communications in Computer and Information Science*. Springer, 2018.

January 2019

Juan Antonio Lossio-Ventura
Hugo Alatrística-Salas

Organization

SIMBig 2018: Organizing Committee

General Organizers

Juan Antonio Lossio-Ventura	University of Florida, USA
Hugo Alatrasta-Salas	Universidad del Pacífico, Peru

Local Organizers

Michelle Rodriguez Serra	Universidad del Pacífico, Peru
Cristhian Ganvini Valcarcel	Universidad Andina del Cusco, Peru
Pilar Hidalgo Leon	Universidad del Pacífico, Peru

SNMAM Track Organizers

Jorge Valverde-Rebaza	Visibilia, Brazil
Alneu de Abdrade Lopes	University of São Paulo, Brazil

DISE Track Organizers

Denisse Muñante Arzapalo	Fondazione Bruno Kessler (FBK), Italy
Nelly Condori Fernandez	Universidade da Coruna, Spain
Carlos Gavidia Calderon	University College London, UK

ANLP Track Organizers

Marco Antonio Sobrevilla-Cabezudo	University of São Paulo, Brazil
Félix Arturo Oncevay-Marcos	Pontificia Universidad Católica del Perú, Peru
Armando Fermin Perez	Universidad Nacional Mayor de San Marcos, Peru

PSCBig Track Organizers

Ali Tekeoglu	SUNY Polytechnic Institute, USA
Miguel Nuñez-del-Prado	Universidad del Pacífico, Peru

SIMBig 2018: Program Committee

SIMBig Program Committee

Nathalie Abadie	COGIT IGN, France
Amine Abdaoui	Stack-Labs, France
César Antonio Aguilar	Pontifica Universidad Católica de Chile, Chile
Frank D. J. Aguilar	University of São Paulo, Brazil

Alexandre Alves	Federal University of São Paulo, Brazil
Sophia Ananiadou	NaCTeM - University of Manchester, UK
Erick Antezana	Norwegian University of Science and Technology, Norway
Ghislain Atemezang	Mondeca, France
Jérôme Azé	LIRMM - University of Montpellier, France
Riza Batista-Navarro	NaCTeM - University of Manchester, UK
Patrice Bellot	University of Aix-Marseille, France
César Beltrán	Pontificia Universidad Católica del Perú, Peru
Mohamed Ben Ellefi	LIS - University of Aix-Marseille, France
Lilian Berton	University of São Paulo, Brazil
Jiang Bian	University of Florida, USA
Albert Bifet	Telecom ParisTech, France
Sandra Bringay	LIRMM - Paul Valéry University, France
Jean-Paul Calbimonte Pérez	HES-SO Valais-Wallis, Switzerland
Thierry Charnois	LIPN - Université Paris 13, France
Marc Chaumont	LIRMM - University of Montpellier, France
Diego Collarana Vargas	University of Bonn, Germany
Nelly Condori-Fernandez	Universidade da Coruña, Spain
Madalina Croitoru	LIRMM - University of Montpellier, France
Bruno Cremilleux	GREYC-CNRS, University of Caen Normandy, France
Gabriela Csurka	NAVER LABS Europe, France
Alneu de Andrade Lopes	University of São Paulo, Brazil
Dina Demner-Fushman	National Library of Medicine, USA
Martín Ariel Domínguez	Universidad Nacional de Córdoba, Argentina
Marcos Dominguez	University of São Paulo, Brazil
Brett Drury	National University of Ireland Galway, Ireland
Paula Estrella	National University of Cordoba, Argentina
Thiago Faleiros	University of São Paulo, Brazil
Karina Figueroa Mora	Michoacán University of San Nicolás de Hidalgo, Mexico
Frédéric Flouvat	PPME - University of New Caledonia, New Caledonia
Giancarlo Fortino	University of Calabria, Italy
Mauro Gaio	LIUPPA, University of Pau, France
Matthias Gallé	NAVER LABS Europe, France
Erick Gomez Nieto	San Pablo Catholic University, Peru
Natalia Grabar	STL CNRS University of Lille 3, France
César Guevara	Complutense University of Madrid, Spain
Jorge Luis Guevara Diaz	Natural Resources Analytics Group IBM Research Lab, Brazil
Adrien Guille	Université Lumière Lyon 2, France
Thomas Guyet	IRISA/LACODAM Agrocampus Ouest, France
Phan Nhat Hai	New Jersey Institute of Technology, USA
Hakim Hacid	Zayed University, United Arab Emirates
Jiawei Han	University of Illinois at Urbana-Champaign, USA
Frank van Harmelen	VU University Amsterdam, The Netherlands

Amanda Hicks	University of Florida, USA
William Hogan	University of Florida, USA
Ian Horrocks	Oxford University, UK
Diana Inkpen	University of Ottawa, Canada
Clement Jonquet	LIRMM - University of Montpellier, France
Alípio Jorge	University of Porto, Portugal
Georgios Kontonatsios	Edge Hill University, UK
Yannis Korkontzelos	Edge Hill University, UK
Ravi Kumar	Google, USA
Nikolaos Lagos	NAVER Labs, France
Juan Guillermo Lazo Lazo	Universidad del Pacífico, Peru
Ulf Leser	Humboldt University of Berlin, Germany
Jose Leomar Todesco	Universidade Federal de Santa Catarina, Brazil
Christian Libaque-Saenz	Universidad del Pacífico, Peru
Cédric López	Emvista, France
Franco Luque	Universidad Nacional de Córdoba and CONICET, Argentina
Maysa Macedo	IBM Research, Brazil
Sabrine Mallek	Higher Institute of Management, France
Florent Massegliaó	Zenith - Inria, France
Héctor Andrés Melgar Sasieta	Pontificia Universidad Católica del Perú, Peru
Claudio Miceli	Federal University of Rio de Janeiro, Brazil
André Miralles	SISO - Irstea, France
Giovanni Montana	University of Warwick, UK
Nils Murrugarra-Llerena	University of Pittsburgh, USA
Mark A. Musen	Stanford University, USA
Jordi Nin	BBVA Data & Analytics and Universidad de Barcelona, Spain
Miguel Nuñez-del-Prado-Cortéz	Universidad del Pacífico, Peru
José Eduardo Ochoa Luna	Catholic University San Pablo, Peru
Maciej Ogródniczuk	Institute of Computer Science, Polish Academy of Sciences, Poland
Jonice Oliveira	Federal University of Rio de Janeiro, Brazil
Thomas Opitz	INRA, France
José Manuel Perea-Ortega	University of Extremadura, Spain
Jessica Pinaire	LIRMM - University of Montpellier, France
Yoann Pitarch	IRIT, France
Bianca Pereira	National University of Ireland, Galway, Ireland
Jorge Poco	San Pablo Catholic University, Peru
Pascal Poncelet	LIRMM - University of Montpellier, France
Mattia Proserpi	University of Florida, USA
Marcos Quiles	Federal University of São Paulo, Brazil
Julien Rabatel	Catholic University of Leuven, Belgium
José-Luis Redondo-García	Amazon, UK

Mathieu Roche	Cirad - TETIS, France
Nancy Rodriguez	LIRMM, CNRS - University of Montpellier, France
Fatiha Saïs	University of Paris-Sud 11, France
Arnaud Sallaberry	LIRMM - Paul Valéry University, France
Nazha Selmaoui-Folcher	PPME - University of New Caledonia, New Caledonia
Matthew Shardlow	University of Manchester, UK
Selja Seppälä	University College Cork, Ireland
Gerardo Sierra-Martínez	Universidad Autónoma de México, Mexico
Newton Spolaor	State University of Western Paraná, Brazil
Claude Tadonki	MINES ParisTech - PSL Research University, France
Alvaro Talavera López	Universidad del Pacífico, Peru
Andon Tchechmedjiev	LIRMM - University of Montpellier, France
Maguelonne Teisseire	Irstea, TETIS, France
Paul Thompson	University of Manchester, UK
Thibaut Thonet	University of Grenoble Alpes, France
Camilo Thorne	Institut für Maschinelle Sprachverarbeitung, University of Stuttgart, Germany
Ilaria Tiddi	Open University, UK
Juan Manuel Torres	University of Avignon, France
Turki Turki	King Abdulaziz University, Saudi Arabia
Willy Ugarte	University of Applied Sciences, Peru
Carlos Vázquez	École de technologie supérieure, Canada
Didier Vega	Universidade de São Paulo, Brazil
Julien Velcin	ERIC Lab - University of Lyon 2, France
Maria-Esther Vidal	Universidad Simón Bolívar, Venezuela
Boris Villazon-Terrazas	Fujitsu Laboratories of Europe, Spain
Sebastian Walter	Semalytix GmbH, Germany
Florence Wang	LIRMM, France
Guo Yi	University of Florida, USA
Osmar Zaiane	University of Alberta, Canada
Amrapali Zaveri	Dumontier Lab, Stanford University, USA
He Zhe	Florida State University, USA
Pierre Zweigenbaum	LIMS-CNRS, France

SNMAM Program Committee

Nazli Bagherzadeh Karimi	National University of Ireland Galway, Ireland
Lilian Berton	University of Santa Catarina State, Brazil
Ricardo Campos	Polytechnic Institute of Tomar and LIAAD/INESC TEC, Portugal
Alexandre Donizeti	Federal University of ABC, Brazil
Brett Drury	Scicrop, Brazil
Thiago de Paulo Faleiros	University of Brasilia, Brazil
Huei Diana Lee	Western Paraná State University, Brazil
Sabrina Mallek	Institut Supérieur de Gestion de Tunis, Tunisia
Ricardo Marcacini	Federal University of Mato Grosso do Sul, Brazil

Nils Murrugarra-Llerena	University of Pittsburgh, USA
Fabricio Olivetti de França	Federal University of ABC, Brazil
Pascal Poncelet	LIRMM - University of Montpellier, France
Ronaldo C. Prati	Federal University of ABC, Brazil
Mathieu Roche	Cirad - TETIS, France
Luca Rossi	Aston University, UK
Rafael Santos	National Institute for Space Research (INPE), Brazil
Pedro Shiguihara Juárez	Universidad Peruana de Ciencias Aplicadas, Peru
Newton Spolaor	State University of Western Paraná, Brazil
Victor Stroele	Federal University of Juiz de Fora, Brazil

ANLP Program Committee

Fernando Emilio Alva Manchego	University of Sheffield, UK
Fernando Antônio Asevedo Nóbrega	University of São Paulo, Brazil
Leandro Borges dos Santos	University of São Paulo, Brazil
Shay Cohen	University of Edinburgh, UK
Paula Christina Figueira Cardoso	Federal University of Lavras, Brazil
Nathan Siegle Hartmman	University of São Paulo, Brazil
Roque Enrique López Condori	Institute for Research in Computer Science and Automation, France
Shashi Narayan	University of Edinburgh, UK
Thiago Alexandre Salgueiro Pardo	University of São Paulo, Brazil
Márcio de Souza Dias	Federal University of Goiás, Brazil
Francis M. Tyers	UiT Norgga árktaš universitehta, Norway

DISE Program Committee

Nour Ali	Brunel University London, UK
Joao Araujo	Universidade Nova de Lisboa, Portugal
Yudith Cardinale	Universidad Simon Bolivar, Venezuela
Alejandro Catala	University of Twente, The Netherlands
Vanea Chiprianov	Université de Pau et des Pays de l'Adour, France
Davide Fucci	University of Hamburg, Germany
David Gomez-Jauregui	ESTIA, France
Vincent Lalanne	Université de Pau et des Pays de l'Adour, France
Cristian Lopez	Universidad de la Salle, Peru
Nazim Madhavji	University of Western Ontario, Canada
Itzel Morales Ramirez	Infotec, Mexico
Manuel Munier	Université de Pau et des Pays de l'Adour, France
Oscar Pastor	Universitat Politècnica de València, Spain
Dietmar Pfahl	University of Tartu, Estonia

Jose Antonio Pow Sang	Pontificia Universidad Católica del Perú, Peru
Daniel Rodriguez	Universidad de Alcalá, Spain
Lizeth Tapia	University of Oslo, Norway
Yuanyuan Zhang	University College London, UK

PSCBig Program Committee

Nihat Altıparmak	University of Louisville, USA
Hisham Kholidy	SUNY Polytechnic Institute, USA
John Marsh	SUNY Polytechnic Institute, USA
Ali Tosun	University of Texas at San Antonio, USA

SIMBig 2018: Organizing Institutions and Sponsors

Organizing Institutions

Universidad del Pacífico, Peru¹
University of Florida, USA²

Collaborating Institutions

Springer, Germany³
Universidad Andina del Cusco, Peru⁴

Sponsoring Institutions

North American Chapter of the ACL, USA⁵
Telefónica del Perú, Peru⁶
BBVA, Peru⁷

SNMAM Organizing Institutions

Visibilia, Brazil⁸
Laboratório de Inteligência Computacional, ICMC, USP, Brazil⁹

¹ <http://www.up.edu.pe/>.

² <http://www.ufl.edu/>.

³ <http://www.springer.com/la/>.

⁴ <http://www.uandina.edu.pe/>.

⁵ <http://naacl.org>.

⁶ <https://www.telefonica.com.pe/>.

⁷ <https://www.bbvacontinental.pe>.

⁸ <http://visibilia.net.br>.

⁹ <http://labic.icmc.usp.br/>.

ANLP Organizing Institutions

Universidad Nacional Mayor de San Marcos, Peru¹⁰

IA Labs, Pontificia Universidad Católica del Perú, Peru¹¹

Laboratório de Inteligência Computacional, ICMC, USP, Brazil¹²

DISE Organizing Institutions

Fondazione Bruno Kessler, Italy¹³

Universidade da Coruna, Spain¹⁴

University College London, UK¹⁵

PSCBig Organizing Institutions

SUNY Polytechnic Institute, USA¹⁶

Universidad del Pacifico, Peru¹⁷



¹⁰ <http://www.unmsm.edu.pe/>.

¹¹ <http://ia.inf.pucp.edu.pe>.

¹² <http://labic.icmc.usp.br/>.

¹³ <https://www.fbk.eu/en/>.

¹⁴ <https://www.udc.es>.

¹⁵ <https://www.ucl.ac.uk>.

¹⁶ <https://sunypoly.edu>.

¹⁷ <http://www.up.edu.pe/>.

Contents

Clinical, Consumer Health, and Visual Question Answering	1
<i>Dina Demner-Fushman</i>	
Which Is the Tallest Building in Europe? Representing and Reasoning About Knowledge	7
<i>Ian Horrocks</i>	
Data-Driven Requirements Engineering. The SUPERSEDE Way.	13
<i>Anna Perini</i>	
Word Embeddings and Deep Learning for Spanish Twitter Sentiment Analysis	19
<i>José Ochoa-Luna and Disraeli Ari</i>	
Twitter Event Detection in a City	32
<i>Martín Steglich, Raúl Speroni, and Juan José Prada</i>	
ANEW for Spanish Twitter Sentiment Analysis Using Instance-Based Multi-label Learning Algorithms	46
<i>Rodrigo Palomino, Carlos Meléndez, David Mauricio, and Jorge Valverde-Rebaza</i>	
An Operational Deep Learning Pipeline for Classifying Life Events from Individual Tweets	54
<i>Xinsong Du, Jiang Bian, and Mattia Prosperi</i>	
Using Behavior and Text Analysis to Detect Propagandists and Misinformers on Twitter	67
<i>Michael Orlov and Marina Litvak</i>	
Analyzing the Retweeting Behavior of Influencers to Predict Popular Tweets, with and Without Considering their Content	75
<i>Matías Gastón Silva, Martín Ariel Domínguez, and Pablo Gabriel Celayes</i>	
Link Prediction in Co-authorship Networks Using Scopus Data	91
<i>Erik Medina-Acuña, Pedro Shiguihara-Juárez, and Nils Murrugarra-Llerena</i>	
Aerial Scene Classification and Information Retrieval via Fast Kernel Based Fuzzy C-Means Clustering	98
<i>Zhengmao Ye, Hang Yin, and Yongmao Ye</i>	

A Case Study of Library Data Management: A New Method to Analyze Borrowing Behavior	112
<i>Luis Cano, Erick Hein, Mauricio Rada-Orellana, and Claudio Ortega</i>	
Sparkmach: A Distributed Data Processing System Based on Automated Machine Learning for Big Data	121
<i>Gussepe Bravo-Rocca, Piero Torres-Robatty, and Jose Fiestas-Iquiria</i>	
Deep Dive into Authorship Verification of Email Messages with Convolutional Neural Network	129
<i>Marina Litvak</i>	
Monitoring of Air Quality with Low-Cost Electrochemical Sensors and the Use of Artificial Neural Networks for the Atmospheric Pollutants Concentration Levels Prediction	137
<i>Ana Luna, Alvaro Talavera, Hector Navarro, and Luis Cano</i>	
Data Mining Algorithms for Risk Detection in Bank Loans	151
<i>Alvaro Talavera, Luis Cano, David Paredes, and Mario Chong</i>	
DETECTOR: Automatic Detection System for Terrorist Attack Trajectories . . .	160
<i>Isaias Hoyos, Bruno Esposito, and Miguel Nunez-del-Prado</i>	
Car Monitoring System in Apartments' Garages by Small Autonomous Car Using Deep Learning.	174
<i>Leonardo León-Vera and Felipe Moreno-Vera</i>	
A Framework for Analytical Approaches to Combine Interpretable Models. . . .	182
<i>Pedro Strecht, João Mendes-Moreira, and Carlos Soares</i>	
Processing Quechua and Guarani Historical Texts Query Expansion at Character and Word Level for Information Retrieval	198
<i>Johanna Cordova, Capucine Boidin, César Itier, Marie-Anne Moreaux, and Damien Nouvel</i>	
Topic Modeling Applied to Business Research: A Latent Dirichlet Allocation (LDA)-Based Classification for Organization Studies	212
<i>Carlos Vilchez-Román, Farita Huamán-Delgado, and Sol Sanguinetti-Cordero</i>	
Using Neural Network for Identifying Clickbaits in Online News Media	220
<i>Amin Omidvar, Hui Jiang, and Aijun An</i>	
Spanish Named Entity Recognition in the Biomedical Domain	233
<i>Viviana Cotik, Horacio Rodríguez, and Jorge Vivaldi</i>	
Ontology Modeling of the Estonian Traffic Act for Self-driving Buses	249
<i>Alberto Nogales, Ermo Täks, and Kuldar Taveter</i>	

Thought Off-line Sanitization Methods for Bank Transactions.	257
<i>Isaias Hoyos and Miguel Nunez-del-Prado</i>	
Big Data for Development: An Approach as a State Government Capacity in the Countries.	265
<i>Marcelino Villaverde Aguilar</i>	
Towards Real-Time Automatic Stress Detection for Office Workplaces	273
<i>Franci Suni Lopez, Nelly Condori-Fernandez, and Alejandro Catala</i>	
SoTesTeR: Software Testing Techniques' Recommender System Using a Collaborative Approach	289
<i>Ronald Ibarra and Glen Rodriguez</i>	
Crowdsourcing High-Quality Structured Data	304
<i>Harry Halpin and Ioanna Lykourantzou</i>	
Ethical and Socially-Aware Data Labels.	320
<i>Elena Beretta, Antonio Vetrò, Bruno Lepri, and Juan Carlos De Martin</i>	
Shadow Removal in High-Resolution Satellite Images Using Conditional Generative Adversarial Networks	328
<i>Giorgio Morales, Samuel G. Huamán, and Joel Telles</i>	
A Mixed Model Based on Shape Context and Spark for Sketch Based Image Retrieval	341
<i>Willy Puenternan Fernández and César A. Beltrán Castañón</i>	
Continuous Detection of Abnormal Heartbeats from ECG Using Online Outlier Detection	349
<i>Yuhang Lin, Byung Suk Lee, and Daniel Lustgarten</i>	
Do Public and Government Think Similar About Indian Cleanliness Campaign?	367
<i>Aarzo Dhiman and Durga Toshniwal</i>	
Author Index	381