Communications in Computer and Information Science 795

Commenced Publication in 2007
Founding and Former Series Editors:
Alfredo Cuzzocrea, Xiaoyong Du, Orhun Kara, Ting Liu, Dominik Ślęzak,
and Xiaokang Yang

Editorial Board

Juan Antonio Lossio-Ventura
Hugo Alatrista-Salas (Eds.)

Information Management and Big Data

4th Annual International Symposium, SIMBig 2017
Lima, Peru, September 4–6, 2017
Revised Selected Papers

Editors
Juan Antonio Lossio-Ventura
University of Florida
Gainesville, FL
USA

Hugo Alatrista-Salas
Universidad del Pacífico
Lima
Peru

Preface

Today, data scientists use the term "big data" to describe the exponential growth and availability of data, which could be structured and unstructured. In this context, techniques used in data science must face a new challenge, which is to extract insights from a large amount of real-time and heterogeneous data (*e.g.*, video, audio, text, image).

Big data has taken place over the past 20 years. For instance, social networks such as Facebook, Twitter, and LinkedIn generate masses of data, which are available to be accessed by other applications. Several domains, including biomedicine, life sciences, and scientific research, have been affected by big data. Therefore there is a need to understand and exploit these data. This process is performed with data science, which is based on methodologies of data mining, natural language processing, Semantic Web, statistics, etc. This allows us to gain new insight through data-driven research. A major problem hampering big data analytics development is the need to process several types of data, such as structured, numeric, and unstructured data (*e.g.*, video, audio, text, image, etc.).

The Annual International Symposium on Information Management and Big Data seeks to present new methods in fields related to the data science for analyzing and managing large volumes of data. SIMBig aims to bring together main — national and international — actors in the field dealing with new technologies dedicated to handling a large amount of information. Moreover, the symposium is a convivial place where these actors present their scientific contributions in the form of full and short papers. This book offers extended versions of the best papers presented at SIMBig 2017[1]. This fourth edition of SIMBig was held in Lima, Peru, during September 4–6. The proceedings are indexed in DBLP[2] [1] and as CEUR Workshop Proceedings[3].

In this special edition, ten long papers were selected from 24 presented in the conference. SIMBig 2017 received 71 submissions.

SIMBig is positioning itself as one of the most important conferences in South America on issues related to information management and big data.

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 the Semantic Web, but also multilingual text processing, and biomedical NLP.

Topics of interest of SIMBig included: data science, big data, data mining, natural language processing, bio-NLP, text mining, information retrieval, machine learning, the Semantic Web, ontologies, Web mining, knowledge representation and linked open data, social networks, social Web, and Web science, information visualization, OLAP,

---

[1] http://simbig.org/SIMBig2017/.

[2] http://dblp1.uni-trier.de/db/conf/simbig/simbig2017.html.

[3] http://ceur-ws.org/Vol-2029/.

data warehousing, business intelligence, spatiotemporal data, health care, agent-based systems, reasoning and logic, constraints, satisfiability, and search.

SIMBig 2017 was supported mainly by the Universidad del Pacífico, the Pontifical Catholic University of Peru, and the University of Florida.

March 2018   Juan Antonio Lossio-Ventura
                                                         Hugo Alatrista-Salas

Reference

1. Juan Antonio Lossio-Ventura and Hugo Alatrista-Salas (eds.), Proceedings of the 4th Annual International Symposium on Information Management and Big Data, SIMBig 2017, Lima, Peru, September 4–6, 2017. CEUR Workshop Proceedings 2029, CEUR-WS.org 2017.

Organization

SIMBig 2017: Organizing Committee

General Organizers

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

Local Organizers

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

SNMAM Track Organizers

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

ANLP Track Organizers

Marco Antonio Sobrevilla-Cabezudo	University of São Paulo, Brazil
Félix Arturo Oncevay-Marcos	Pontificia Universidad Católica del Perú, Peru
Félix Armando Fermín-Pérez	UNMSM, Peru

SIMBig 2017: Program Committee

SIMBig Program Committee

Elie Abi-Lahoud	University College Cork, Ireland
César Antonio Aguilar	Pontificia Universidad Católica de Chile, Chile
Sophia Ananiadou	NaCTeM University of Manchester, UK
Jérôme Azé	LIRMM University of Montpellier, France
Riza Batista-Navarro	NaCTeM University of Manchester, UK
Nicolas Béchet	IRISA Université de Bretagne-Sud, France
Jiang Bian	University of Florida, USA
Albert Bifet	MINES ParisTech, France
Sandra Bringay	LIRMM Paul Valéry University, France
Bruno Cremilleux	Université de Caen Normandie, CNRS
Fabio Crestani	University of Lugano, Switzerland
Martín Ariel Domínguez	Universidad Nacional de Córdoba, Argentina
Brett Drury	National University of Ireland Galway, Ireland

Frédéric Flouvat	PPME University of New Caledonia, New Caledonia
Philippe Fournier-Viger	Harbin Institute of Technology, China
Natalia Grabar	University of Lille 3, France
Adrien Guille	Université Lumière Lyon 2, France
Thomas Guyet	IRISA/LACODAM Agrocampus Ouest, France
Phan Nhat Hai	New Jersey Institute of Technology, USA
Jiawei Han	University of Illinois, USA
Sébastien Harispe	Ecole des Mines d'Alès, France
William Hogan	University of Florida, USA
Vijay Ingalalli	Inria Bretagne Atlantique, France
Georgios Kontonatsios	Edge Hill University, UK
Yannis Korkontzelos	Edge Hill University, UK
Ravi Kumar	Google, USA
Christian Libaque-Saenz	Universidad del Pacífico, Peru
Cédric López	VISEO Research and Development Unit, France
André Miralles	SISO Team, France
François Modave	University of Florida, USA
Jordi Nin	BBVA Data & Analytics and Universidad de Barcelona, Spain
Miguel Nuñez-del-Prado-Cortéz	Universidad del Pacífico, Peru
Maciej Ogrodniczuk	Polish Academy of Sciences, Poland
Marco Aurélio Pacheco	Pontifícia Universidade Católica do Rio de Janeiro, Brazil
José Manuel Perea-Ortega	University of Extremadura, Spain
Pascal Poncelet	LIRMM University of Montpellier, France
Julien Rabatel	Catholic University of Leuven, Belgium
José-Luis Redondo-García	Polytechnic University of Madrid, Spain
Mathieu Roche	Cirad - TETIS - LIRMM, France
Nancy Rodriguez	LIRMM University of Montpellier, France
Rafael Rossi	University of São Paulo, Brazil
Fatiha Saïs	Université Paris-Sud 11, France
Arnaud Sallaberry	LIRMM Paul Valéry University, France
Matthew Shardlow	University of Manchester, UK
Gerardo Eugenio Sierra-Martínez	Instituto de Ingeniería, UNAM
Newton Spolaor	Universidade de São Paulo, Brazil
Claude Tadonki	MINES ParisTech, France
Maguelonne Teisseire	Irstea, TETIS, France
Paul Thompson	University of Manchester, UK
Carlos Vàzquez	École de technologie supérieure, Canada
Didier Vega	Universidade de São Paulo, Brazil
Julien Velcin	Université Lumière Lyon 2, France
Maria-Esther Vidal	Universidad Simón Bolívar, Venezuela
Boris Villazon-Terrazas	Fujitsu Laboratories of Europe, Spain

Youyou Wu	Kellogg School of Management, USA
Yang Yang	Kellogg School of Management, Northwestern University, USA
Guo Yi	University of Florida, USA
Amrapali Zaveri	Dumontier Lab, USA
He Zhe	Florida State University, USA

SNMAM Program Committee

Alan Valejo	University of São Paulo, Brazil
Brett Drury	National University of Ireland Galway, Ireland
Celso Kaestner	Federal University of Technology of Paraná, Brazil
Didier Vega-Oliveros	University of São Paulo, Brazil
Hugo Gualdron Colmenares	University of São Paulo, Brazil
José Benito Camiña	Tecnológico de Monterrey, Mexico
Jesús Mena-Chalco	Federal University of ABC, Brazil
Lilian Berton	University of Santa Catarina State, Brazil
Luca Rossi	Aston University, UK
Marcos Domingues	State University of Maringá, Brazil
Marcos G. Quiles	Federal University of São Paulo, Brazil
Mathieu Roche	CIRAD and University of Montpellier, France
Merley Conrado	Intel Corp., USA
Newton Spolaôr	Western Paraná State University, Brazil
Pascal Poncelet	University of Montpellier, France
Rafael Rossi	Federal University of Mato Grosso do Sul, Brazil
Ricardo Campos	Polytechnic Institute of Tomar and LIAAD/INESC TEC, Portugal
Ricardo Marcacini	Federal University of Mato Grosso do Sul, Brazil
Ronaldo C. Prati	Federal University of ABC, Brazil
Sabrine Mallek	Institut Supérieur de Gestion de Tunis, Tunisia
Thiago de Paulo Faleiros	University of Brasilia, Brazil
Vânia Neves	Federal University of Juiz de Fora, Brazil
Victor Stroele	Federal University of Juiz de Fora, Brazil

ANLP Program Committee

Thiago Alexandre Salgueiro Pardo	University of São Paulo, Brazil
Nathan Siegle Hartmman	University of São Paulo, Brazil
Leandro Borges dos Santos	University of São Paulo, Brazil
Fernando Emilio Alva Manchego	University of Sheffield, UK
Paula Christina Figueira Cardoso	Federal University of Lavras, Brazil
Márcio de Souza Dias	Federal University of Goiás, Brazil
Fernando Antônio Asevedo Nóbrega	University of São Paulo, Brazil

Roque Enrique López   Institute for Research in Computer Science
    Condori   and Automation, France
Francis M. Tyers   UiT Norgga árktalaš universitehta, Norway
Shay Cohen   University of Edinburgh, UK
Shashi Narayan   University of Edinburgh, UK

SIMBig 2017: Organizing Institutions and Sponsors

Organizing Institutions

Universidad del Pacífico, Perú[1]
University of Florida, USA[2]
Universidad Andina del Cusco, Perú[3]

Collaborating Institutions

Springer[4]
Banco de Crédito del Perú[5]
Escuela de Post-grado de la Pontificia Universidad Católica del Perú[6]

SNMAM Organizing Institutions

Instituto de Ciências Matemáticas e de Computação, USP, Brazil[7]
Labóratorio de Intêligencia Computacional, ICMC, USP, Brazil[8]
Universidade Federal de São Carlos, Brazil[9]

ANLP Organizing Institutions

Universidad Nacional Mayor de San Marcos, Perú[10]
Grupo de Reconocimiento de Patrones e Inteligencia Artificial Aplicada, PUCP, Perú[11]
Instituto de Ciências Matemáticas e de Computação, USP, Brazil
Universidade Federal de São Carlos, Brazil

---

[1] http://www.up.edu.pe/.

[2] http://www.ufl.edu/.

[3] http://www.uandina.edu.pe/.

[4] http://www.springer.com/la/.

[5] https://www.viabcp.com/wps/portal/.

[6] http://posgrado.pucp.edu.pe/la-escuela/presentacion/.

[7] http://www.icmc.usp.br/Portal/.

[8] http://labic.icmc.usp.br/.

[9] http://www2.ufscar.br/home/index.php.

[10] http://www.unmsm.edu.pe/.

[11] http://inform.pucp.edu.pe/∼grpiaa/.

Contents