

List of contributions

1	Overview of SIMBig 2015	10
·	Overview of SIMBig 2015: 2nd Annual International Symposium on Information Management and Big Data <i>Juan Antonio Lossio-Ventura and Hugo Alatrasta-Salas</i>	10
2	Keynote Speaker Presentations	13
·	Real-Time Big Data Stream Analytics <i>Albert Bifet</i>	13
·	Detecting Locations from Twitter Messages <i>Diana Inkpen</i>	15
·	Opinion Mining: Taking into account the Criteria! <i>Pascal Poncelet</i>	17
·	How to Deal with Heterogeneous Data? <i>Mathieu Roche</i>	19
·	Rich Data: Risks, Issues, Controversies & Hype <i>Osmar R. Zaiane</i>	21
3	Session 1: Text Mining and Social Networks	25
·	Dependency-based Topic-Oriented Sentiment Analysis in Microposts <i>Prasadith Buddhitha and Diana Inkpen</i>	25
·	Specializations for the Peruvian Professional in Statistics: A Text Mining Approach <i>Luis Angel Cajachahua Espinoza, Andrea Ruiz Guerrero and Tomás Nieto Agudo</i>	35
·	Social Networks of Teachers on Twitter <i>Hernán Gil Ramírez and Rosa María Guilleumas</i>	43
·	Esopo: Sensors and Social Pollution Measurements <i>Vittoria Cozza, Ilaria Guagliardi, Michelangelo Rubino, Raffaele Cozza, Alessandra Martello, Marco Picelli and Eustrat Zhupa</i>	52
4	Session 2: Special Track - WTI	58
·	Causation Generalization Through the Identification of Equivalent Nodes in Causal Sparse Graphs Constructed from Text using Node Similarity Strategies <i>Brett Drury, Jorge Valverde-Rebaza and Alneu de Andrade Lopes</i> . . .	58
·	Text Mining Applied to SQL Queries: A Case Study for the SDSS SkyServer <i>Vitor Hirota Makiyama, M. Jordan Raddick, Rafael Santos</i>	66
·	Spreader Selection by Community to Maximize Information Diffusion in Social Networks <i>Didier Vega-Oliveros and Lilian Berton</i>	73
·	Behavior of Symptoms on Twitter <i>Dennis Salcedo and Alejandro León</i>	83

5	Session 3: Biomedical Informatics	85
·	Scalability Potential of BWA DNA Mapping Algorithm on Apache Spark <i>Zaid Alars and Hamid Mushtaq</i>	85
·	Explain Sentiments using Conditional Random Field and a Huge Lexical Network <i>Mike Donald Tapi-Nzali, Sandra Bringay, Pierre Pompidor, Joël Maïzi, Christian Lavergne and Caroline Mollevi</i>	89
6	Session 4: Databases	94
·	A Security Price Data Cleaning Technique: Reynold's Decomposition Approach <i>Rachel Mok, Wai Mok and Kit Cheung</i>	94
·	Automation of Process to Load Database from OSM for the Design of Public Routes <i>Gissella Bejarano, José Astuvilca and Pedro Vega</i>	99
7	Session 5: Big Data	106
·	Big Data Architecture for Prediction of Number Portability in Mobile Phone Companies <i>Alonso Raúl Melgarejo Galván and Katerine Rocío Clavo Navarro</i>	106
·	MapReduce and Relational Database Management Systems: Competing or Completing Paradigms? <i>Dhouha Jemal and Rim Faiz</i>	117
·	Performance of Alternating Least Squares in a Distributed Approach Using GraphLab and MapReduce <i>Laura Cruz, Elizabeth Veronica Vera Cervantes and José Eduardo Ochoa Luna</i>	122
·	Data Modeling for NoSQL Document-Oriented Databases <i>Harley Vera Olivera, Maristela Holanda, Valeria Guimarães, Fernanda Hondo and Wagner Boaventura</i>	129
8	Session 6: Posters	136
·	Hipi, as Alternative for Satellite Images Processing <i>Wilder Nina Choquehuayta, René Cruz Muñoz, Juber Serrano Cervantes, Alvaro Mamani Aliaga, Pablo Yanyachi and Yessenia Yari</i>	136
·	Multi-agent System for Usability Improvement of an University Administrative System <i>Felix Armando Fermin Perez and Jorge Leoncio Guerra Guerra</i>	138