# List of contributions


**1 SIMBig 2016 Overview** **10**

· Information Management and Big Data: SIMBig overview
*Juan Antonio Lossio-Ventura, Hugo Alatrista-Salas* . . . . . 10

**2 Keynote summaries** **15**

· Mining Internet of Things (IoT) Big Data Streams
*Albert Bifet* . . . . . 15

· Spatio-Temporal Data Mining: From Big Data to Patterns
*Maguelonne Teisseire* . . . . . 17

· Personalised Recommendations for Context Aware Suggestions
*Fabio Crestani* . . . . . 19

**3 Full papers** **22**

· Construction of a biodiversity knowledge repository using a text mining-based framework
*Riza Batista-Navarro, Chrysoula Zerva and Sophia Ananiadou* . . . . . 22

· Measuring citizen participation in South African public debates using Twitter: An exploratory study
*Selvas Mwanza and Hussein Suleman* . . . . . 26

· Opinion Analysis Applied to Politics: A case study based on Twitter
*Gilberto Nunes, Denivaldo Lopes and Zair Abdelouahab* . . . . . 35

· Predicting Reactions to Blog Headlines
*Rel Guzman, Jose Eduardo Ochoa-Luna, Laura Cruz-Quispe and Elizabeth Vera-Cervantes* . . . . . 43

· A Corpus Builder: Retrieving Raw Data from GitHub for Knowledge Reuse In Requirements Elicitation
*Roxana Lisette Quintanilla Portugal, Hugo Roque, Julio Cesar Sampaio do Prado Leite* . . . . . 48

· Cloud risk communication in social media: The case of Premera Blue Cross
*Jean Pierre Guy Gashami, Christian Fernando Libaque-Saenz, Myeong-Cheol Park and Jae Jeung Rho* . . . . . 55

· The impact of network sampling on relational classification
*Lilian Berton, Didier A. Vega-Oliveros, Jorge Valverde-Rebaza, Andre Tavares Da Silva and Alneu de Andrade Lopes* . . . . . 62

· Measuring more than just exchange: Are multiplex networks the key to understanding informal economic relationships?
*John Harvey* . . . . . 73

· Dictionary-Based Sentiment Analysis applied to specific domain using Web Mining approach
*Laura Cruz, Jose Ochoa, Mathieu Roche and Pascal Poncelet* . . . . . 80

· An Empirical Method Exploring a Large Set of Features for Authorship Identification
*Seifeddine Mechti, Rim Faiz, Maher Jaoua and Lamia Hadrich Belguith* 89

· In God we trust. All others must bring data. –W. Edwards Deming.
Using word embeddings to recognize idioms
*Jing Peng and Anna Feldman* . . . . . 96
· Evaluating Class Assignment Semantic Redundancy on Linked Datasets
*Leandro Mendoza and Alicia Diaz* . . . . . 103
· Knowledge Tier Platform for Graph Mining in (Smart) Cities
*Miguel Nunez-Del-Prado, Edgardo Bravo, Miguel Sierra, Isaias Hoyos and Miguel Canchay* . . . . . 110
· Social Spider Algorithm Approach for Clustering
*Harley Vera-Olivera, José Luis Soncco-Álvarez and Lauro Enciso-Rodas* 114
· A clustering optimization approach for disaster relief delivery: A case study in Lima-Peru
*Jorge Vargas-Florez, Rosario Medina-Rodriguez and Rafael Alva-Cabrera* 122
· DAZIO: Detecting Activity Zones based on Input/Output call and SMS activity
*Miguel Nunez-Del-Prado, Ana Luna and Romain Gauthier* . . . . . 130
· Sense-Level Semantic Clustering of Hashtags in Social Media
*Ali Javed and Byung Suk Lee* . . . . . 140
· Russian Named Entities Recognition and Classification Using Distributed Word and Phrase Representations
*Roman Ivanitsky, Alexander Shipilo and Liubov Kovriguina* . . . . . 150
· Big Data Analytics for Opinion Mining and Patterns Detection of the Tunisian Election
*Zeineb Dhouioui, Hanen Bouali and Jalel Akaichi* . . . . . 157
· Use of text mining for understanding Peruvian students and faculties' perceptions on bibliometrics training
*Carlos Vílchez-Román and Joel Alhuay-Quispe* . . . . . 165