



# **ALLDATA 2018**

The Fourth International Conference on Big Data, Small Data, Linked Data and  
Open Data

## **KESA 2018**

The International Workshop on Knowledge Extraction and Semantic Annotation

April 22 - 26, 2018

Athens, Greece

## **ALLDATA 2018 Editors**

Gary Weckman, Ohio University, USA

Jerzy Grzymala-Busse, University of Kansas, USA

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

Printed by Curran Associates, Inc. (423: )

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

Small Data: Applications and Architecture <i>Cheng-Kang Hsieh, Faisal Alquaddoomi, Fabian Okeke, John P. Pollak, Lucky Gunasekara, and Deborah Estrin</i>	1
Knowledge Base L-V-C Mapping Method <i>Dong-Jae Lee, Yun-Hee Son, and Kyu-Chul Lee</i>	11
Data Provenance Service Prototype for Collaborative Data Infrastructures <i>Vasily Bunakov, Javier Quinteros, and Linda Reijnhoudt</i>	15
Design and Implementation of Candlestick Chart Retrieval Algorithm for Predicting Stock Price Trend <i>Yoshihisa Udagawa</i>	19
Design of Elastic Hadoop Supporting Dynamic Scaling of the Cluster <i>Wooseok Ryu</i>	26
A New Representation of Air Traffic Data Adapted to Complexity Assessment <i>Georges Mykoniatis, Florence Nicol, and Stephane Puechmorel</i>	28
Monitoring of Coastal Environments Using Data Mining <i>Corneliu Octavian Dumitru, Gottfried Schwarz, and Mihai Datcu</i>	34
Descriptive Sentence Extraction for Text to 3D Scene Generation <i>Valentina Bova and Elena Cardillo</i>	40
Application of Event Sourcing in Research Data Management <i>Jedrzej Rybicki</i>	46
A Big Data Quality Preprocessing and Domain Analysis Provisioner Framework Using Cloud Infrastructures <i>Dirk Holscher, Timo Bayer, Philipp Ruf, Christoph Reich, and Frank Gut</i>	53
PKGAWAS: A Knowledge Services and Allergy Early Warning System of Pollinosis Based on Cross-Border Data Integration <i>Xiao Lei Xiu, Si Zhu Wu, Qing Qian, Jia Wei Cui, and Xiao Kang Sun</i>	59
A Data Clustering Approach for Automated Optical Inspection of Metal Work Pieces <i>Ruth Tesfaye Zibello, Stephan Trahasch, and Tobias Lauer</i>	64
Optimizing Mixed Fuzzy-Rule Formation by Controlled Evolutionary Strategy <i>Matthias Lerner, Hendrik Kuijs, and Christoph Reich</i>	69

Wine Critic Scores and Consumer Behavior in a Major USA Metropolitan Market <i>Andrew Snow and Gary Weckman</i>	75
--	----

On the Number of Conditions in Mining Incomplete Data Using Characteristic Sets and Maximal Consistent Blocks <i>Patrick G. Clark, Cheng Gao, Jerzy W. Grzymala-Busse, and Teresa Mroczek</i>	84
--	----