ALLDATA 2016

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

KESA 2016

The International Workshop on Knowledge Extraction and Semantic Annotation

February 21 - 25, 2016

Lisbon, Portugal

ALLDATA 2016 Editors

Venkat Gudivada, East Carolina University, USA
Dumitru Roman, SINDEF/University of Oslo, Norway
Maria Pia di Buono, University of Salerno, Italy
Mario Monteleone, University of Salerno, Italy

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© (2016) by International Academy, Research, and Industry Association (IARIA)
Please refer to the Copyright Information page.

Printed by Curran Associates, Inc. (2016)

International Academy, Research, and Industry Association (IARIA)
412 Derby Way
Wilmington, DE 19810

Phone: (408) 893-6407
Fax: (408) 527-6351

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: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

AScale: Simple and Fast ETL+Q Scaling for Small and Big Data 1
Pedro Martins, Maryam Abbasi, and Pedro Furtado

Informed Consent and Privacy of De-Identified Information and Estimated Data: Lessons from Iceland and the United States in an Era of Computational Genomics 7
Donna M. Gitter

Social Sentiment Indices Powered by X-Scores 12
Brian Davis, Keith Cortis, Laurentiu Vasiliu, Adamantios Koumpis, Ross McDermott, and Siegfried Handschuh

Stream Mining Revised 18
Sayaka Akioka

Data Quality Centric Application Framework for Big Data 24
Venkat Gudivada, Dhana Rao, and William Grosky

Combining “Small Data” from Surveys and “Big Data” from Online Experiments at Pinterest 33
Jolie Martin

Unsupervised Aircraft Trajectories Clustering: a Minimum Entropy Approach 35
Florence Nicol and Stephane Puechmorel

Modelling Support for a Linked Data Approach to Tool Interoperability 42
Jad El-khoury, Didem Gurdur, Frederic Loiret, Martin Torngren, Da Zhang, and Mattias Nyberg

Reasoning with Place Information on the Linked Data Web 48
Alia Abdelmoty and Khalid Al-Muzaini

Graph-based Data Integration in EUDAT Data Infrastructure 54
Jedrzej Rybicki, Vasily Bunakov, Paolo D'Onorio De Meo, Stephan Kindermann, and Anna Queralt

IPR Issues in Data Sharing via Linkage of Platforms and Apps 59
Iryna Lishchuk, Marc Stauch, and Nikolaus Forgo

Modelling the Cost of Open Data 65
Jolon Faichney, Bela Stantic, Yasaman Moaven, Sanjeev Hiremath, and John Galvin

Link Detection Based on Named Entity Keywords in Turkish News Corpus 71
Hamid Ahmadlouei, Hayri Sever, and Erhan Mengusoglu

Using Domain Ontologies for Classification and Semantic Interpretation of Documents 76
Samia Iltache, Catherine Comparot, Malik Si Mohammed, and Pierre-Jean Charrel

Extraction of Business Information on the web to Supply a Geolocated Search Service 82
Armel Fotsoh Tawofaing, Christian Sallaberry, Annig Le Parc - Lacayrelle, and Tanguy Moal

A Semantic Annotation Model for Syntactic Parsing in Ontology Learning Task 86
Maria Pia di Buono and Mario Monteleone

Semantic Annotation to Support Decision Making 90
Francesca Parisi

An Ontology-based Method for Discovering Specific Concepts from Texts via Knowledge Completion 96
Celine Alec, Chantal Reynaud-Delaitre, and Brigitte Safar

Analyzing the Retirement Satisfaction Predictors among Men and Women Using a Multi-Layer Feed Forward Neural Network and Decision Trees 102
Ehsan Ardjmand, Gary Weckman, Diana Schwerha, and Andrew Snow

Group Method of Data Handling: How Does it Measure Up? 108
Poorani Selvaraj, Gary Weckman, and Andrew Snow

A Novel Event-based Method for Abducting Evolutional Motivation of Empirical Engineering Knowledge 114
Xinyu Li, Zuhua Jiang, Lijun Liu, and Puling Liu