

2015 IEEE International Conference on Information Reuse and Integration (IRI 2015)

**San Francisco, California, USA
13-15 August 2015**



**IEEE Catalog Number: CFP15IRI-POD
ISBN: 978-1-4673-6657-1**

2015 IEEE 16th International Conference on Information Reuse and Integration

IRI 2015

Table of Contents

Foreword.....	xiv
Message from Program Co-Chairs.....	xv
Conference Organizers.....	xvii
International Technical Program Committee.....	xix
Steering Committee.....	xxi
Keynotes.....	xxii
Tutorials.....	xxix
Industry Speakers.....	xxxiii
Sponsors.....	xxxvi

Session A1

Session A11: Semantic and Big Data Management

A Multi-dimensional Comparison of Toolkits for Machine Learning with Big Data	1
<i>Aaron N. Richter, Taghi M. Khoshgoftaar, Sara Landset, and Tawfiq Hasanin</i>	
SDPA: Sensor Data Processing Architecture for Modeling Semantic Data from Sensor Steams	9
<i>Seungmin Seo, Sejin Chun, Byungkook Oh, and Kyong-Ho Lee</i>	
Choosing an Appropriate Ensemble Classifier for Balanced Bioinformatics Data	17
<i>Alireza Fazelpour, Taghi M. Khoshgoftaar, David J. Dittman, and Amri Napolitano</i>	
Real-Time Stream Data Analytics for Multi-purpose Social Media Applications	25
<i>Satyen Abrol, Gunasekar Rajasekar, Latifur Khan, Vaibhav Khadilkar, Siddarth Nagarajan, Nathan McDaniel, Gautam Ganesh, and Bhavani Thuraisingham</i>	

Session A12: Social Media/Networks and Web

What Happens to My Online Social Estate When I Am Gone? An Integrated Approach to Posthumous Online Data Management	31
<i>Leila Bahri, Barbara Carminati, and Elena Ferrari</i>	
Social Media Data Aggregation and Mining for Internet-Scale Customer Relationship Management	39
<i>Stephen Wan, Cecile Paris, and Dimitrios Georgakopoulos</i>	
Smart Maps through Semantic Web, Social Media, and Sentiment Analysis	49
<i>Jose Andres Bolaños Perera, Du Zhang, and Meiliu Lu</i>	
Integration of Heterogeneous Web Services for Event-Based Social Networks	57
<i>Yinuo Zhang, Hao Wu, Anand Panangadan, and Viktor K. Prasanna</i>	

Session A13: Multimedia Data Management

Negative-Based Sampling for Multimedia Retrieval	64
<i>Hsin-Yu Ha, Shu-Ching Chen, and Mei-Ling Shyu</i>	
Perceptually Lossless Image Coding Based on Foveated JND	72
<i>Yiming Li, Hongyi Liu, and Zhenzhong Chen</i>	
Visual and Textual Feature Fusion for Automatic Customs Tariff Classification	76
<i>Bilgehan Turhan, Gozde B. Akar, Cigdem Turhan, and Cihan Yukse</i>	
Ensemble Learning from Imbalanced Data Set for Video Event Detection	82
<i>Yimin Yang and Shu-Ching Chen</i>	

Session A14: Reuse and Integration

The Urban Data Re-use and Integration Platform for Australia: Design, Realisation, and Case Studies	90
<i>Richard O. Sinnott, C. Bayliss, A. Bromage, G. Galang, Y. Gong, P. Greenwood, G. Jayaputera, D. Marques, L. Morandini, M. Nino-Ruiz, G. Nogoarani, H. Pursultani, R. Rabanal, M. Sarwar, W. Voorsluys, and I. Widjaja</i>	
Context and Encapsulation Modeling for Knowledge Reuse	98
<i>Thouraya Bouabana-Tebibel, Stuart H. Rubin, and Asmaa Chebba</i>	
Discovering Reusable Hardware Using Birthmarking Techniques	106
<i>Kevin Zeng and Peter Athanas</i>	
Integrating Open Source Software Repositories on the Web through Linked Data	114
<i>Aftab Iqbal and Stefan Decker</i>	

Session A2

Session A22: Sensing and Sensor Data Management

Multi-layer Sparse Coding Based Ship Detection for Remote Sensing Images	122
<i>Zimeng Li, Daiqin Yang, and Zhenzhong Chen</i>	
Complex Sensor Mashups for Linking Sensors and Formula-Based Knowledge Bases	126
<i>Wonwoo Ro, Giyong Park, Sejin Chun, and Kyong-Ho Lee</i>	
A Novel Small Fault Detection Scheme Based on Local Approach and Divided Difference Filter for Inertial Sensors in Unmanned Aerial Vehicles	134
<i>Chang Liu and Honglun Wang</i>	
Cooperative Multi-sensor Multi-vehicle Localization in Vehicular Adhoc Networks	142
<i>Sepideh Afkhami Goli, Behrouz H. Far, and Abraham O. Fapojuwo</i>	

Session A23: Clouds and Grids

On Verifying Cycle-Freeness in OSGi Applications	150
<i>Ahmed Amar Debza, Chafia Bouanaka, and Nadia Zeghib</i>	
DCCSOA: A Dynamic Cloud Computing Service-Oriented Architecture	158
<i>Mehdi Bahrami. Mukesh Singhal</i>	

Session A24: NLP, Fuzzy Systems, and Networks

PNA: Partial Network Alignment with Generic Stable Matching	166
<i>Jiawei Zhang, Weixiang Shao, Senzhang Wang, Xiangnan Kong, and Philip S. Yu</i>	
Gender Prediction in Random Chat Networks Using Topological Network Structures and Masked Content	174
<i>Michael Crawford and Xingquan Zhu</i>	
Uncertainty Nonlinear Systems Modeling with Fuzzy Equations	182
<i>Raheleh Jafari and Wen Yu</i>	

Poster Session

Sentiment Analysis of Social Network Content	189
<i>Shambhavi Dinakar, Pankaj Andhale, and Manjeet Rege</i>	
XQuery Data Mining Tool for Campus Security	193
<i>Rohit Surve, Meiliu Lu, and Jun Dai</i>	
Using Random Undersampling to Alleviate Class Imbalance on Tweet Sentiment Data	197
<i>Joseph Prusa, Taghi M. Khoshgoftaar, David J. Dittman, and Amri Napolitano</i>	

Topic Discovery and Future Trend Prediction Using Association Analysis and Ensemble Forecasting	203
<i>Jose Hurtado, Shihong Huang,, and Xingquan Zhu</i>	
Enabling Linguistic Analysis of Scientific Metadata through Internationalizing NASA JPL's PODAAC	207
<i>Lewis J. McGibbney, Kim D. Whitehall, Chris A. Mattmann, and Phillip M. Carter</i>	
Accessibility Stable Analysis Pattern (Stable Pattern for Model Based Software Reuse)	211
<i>M.E. Fayad and Siddharth Jindal</i>	
Extending Spark Analytics through Tika-Based Information Extraction and Retrieval	215
<i>Rishi Verma and Chris Mattmann</i>	
A Survey of 2D Face Databases	219
<i>Gabriel Castaneda and Taghi M. Khoshgoftaar</i>	
An Ant Colony Optimization Heuristic to Optimize Prediction of Stability of Object-Oriented Components	225
<i>Haidar Harmanani, Danielle Azar, Grace Zgheib, and David Kozhaya</i>	

Session B1

Session B11: Classification, Clustering, & Mining I

Building an Effective Classification Model for Breast Cancer Patient Response Data	229
<i>Brian Heredia, Taghi M. Khoshgoftaar, Alireza Fazelpour, and David J. Dittman</i>	
A Local SVD Framework for Stable Feature Selection for Clustering	236
<i>Salem Alelyani and Huan Liu</i>	
Simplex Projection for Land Cover Information Mining from Landsat-5 TM Data	244
<i>Uttam Kumar, Cristina Milesi, Sangram Ganguly, S. Kumar Raja, and Ramakrishna R. Nemani</i>	
Using Ensemble Learners to Improve Classifier Performance on Tweet Sentiment Data	252
<i>Joseph Prusa, Taghi M. Khoshgoftaar, and Daivd J. Dittman</i>	

Session B12: Heuristics/Mining & Information Retrieval/Extraction

Gaussian Mixture Model-Based Subspace Modeling for Semantic Concept Retrieval	258
<i>Chao Chen, Mei-Ling Shyu, and Shu-Ching Chen</i>	
On Extracting Arguments in Favor of One Given Conclusion	266
<i>Éric Grégoire</i>	
How Can We Explore Patterns of Customer Segments' Structural Changes? A Sequential Rule Mining Approach	273
<i>Elham Akhond Zadeh Noughabi, Amir Albadvi, and Behrouz Homayoun Far</i>	
City Recorder: Virtual City Tour Using Geo-referenced Videos	281
<i>Guangqiang Zhao, Mingjin Zhang, Tao Li, Shu-ching Chen, and Naphtali Rishe</i>	

Workshop on Formal Methods Integration (FMI 2015)

Formal Methods for Modelling and Analysis of Single-Event Upsets	287
<i>René Rydhof Hansen, Kim Guldstrand Larsen, Mads Chr. Olesen, and Erik Ramsgaard Wognsen</i>	
A Lattice-Based Representation of Temporal Failures	295
<i>André Luis Ribeiro Didier and Alexandre Mota</i>	
Specifying Avionic Embedded Systems by Denotations of the Time-Triggered Constraint-Based Calculus	303
<i>Sardaouna Hamadou, John Mullins, Christophe Chareton, and Abdelouahed Gherbi</i>	
RenewGrass - A Tool for Building Scientific Workflows: Application to the Remote Sensing Domain	311
<i>Sofiane Bendoukha, Hayat Bendoukha, and Daniel Moldt</i>	
Capturing and Verifying Dynamic Program Behaviour Using UML Communication Diagrams and Pi-Calculus	318
<i>Aissam Belghiat, Allaoua Chaoui, and Mokhtar Beldjehem</i>	

Information Reuse and Integration in Health Informatics (IRI-HI 2015)

On Finding Explicit Rules for Personalized Forecasting of Obstructive Sleep Apnea Episodes	326
<i>Ivanoe De Falco, Giuseppe De Pietro, and Giovanna Sannino</i>	
Mastering Situation Awareness in Healthcare Database Systems	334
<i>Andreas Behrend, Philip Schmiegelt, Dieter Gawlick, Zhen Hua Liu, and Ronny Fehling</i>	

Alterations to the Bootstrapping Process within Random Forest: A Case Study on Imbalanced Bioinformatics Data	342
<i>Taghi M. Khoshgoftaar, Alireza Fazelpour, David J. Dittman, and Amri Napolitano</i>	

Session B2

Session B21: Classification, Clustering, & Mining II

A Multi-topic Meta-classification Scheme for Analyzing Lobbying Disclosure Data	349
<i>Xinpeng L. Liao, Chengcui Zhang, Ariel D. Smith, and Grant T. Savage</i>	
Classifying Galaxy Images through Support Vector Machines	357
<i>Kathy Applebaum and Du Zhang</i>	
Fast Text Classification Using Randomized Explicit Semantic Analysis	364
<i>Aibek Musaev, De Wang, Saajan Shridhar, and Calton Pu</i>	
Observing the Effect of the Choice of Classifier on Bioinformatics Data with Varying Levels of Data Quality and Class Balance	372
<i>Alireza Fazelpour, Taghi M. Khoshgoftaar, David J. Dittman, and Ahmad Abu Shanab</i>	

Session B22: Testing, Software Reuse, & Automation

Test Reactive Systems with Buchi Automata: Acceptance Condition Coverage Criteria and Performance Evaluation	380
<i>Bolong Zeng and Li Tan</i>	
Modeling Reusable and Interoperable Faceted Browsing Systems with Category Theory	388
<i>Daniel R. Harris</i>	
On the Need of Integrating Social Media Channels and Open Source Software Repositories	396
<i>Aftab Iqbal and Stefan Decker</i>	

Session C1

Session C11: Information Security and Privacy I

BEAN: A BEhavior ANalysis Approach of URL Spam Filtering in Twitter	403
<i>De Wang and Calton Pu</i>	
On Heuristic Randomization and Reuse as an Enabler of Domain Transference	411
<i>Stuart H. Rubin, Thouraya bouabana-Tebibel, Yasmin Hoadjli, Kadaouia Habib, and Boussaad Yacine Belamiri</i>	

Hierarchical Secure Information and Resource Sharing in OpenStack Community Cloud	419
<i>Yun Zhang, Farhan Patwa, Ravi Sandhu, and Bo Tang</i>	
Privacy-Preserving Protocols for Shortest Path Discovery over Outsourced Encrypted Graph Data	427
<i>Bharath K. Samanthula, Fang-Yu Rao, Elisa Bertino, and Xun Yi</i>	

Session C12: Data Management

DBLPminer: A Tool for Exploring Bibliographic Data	435
<i>Tony Le and Du Zhang</i>	
Datom: Towards Modular Data Management	443
<i>Verena Kantere</i>	
Design and Implementation of a Semantic Web-Based Inference Controller: A Summary	451
<i>Bhavani Thuraisingham, Tyrone Cadenhead, Murat Kantarcioglu, and Vaibhav Khadilkar</i>	
The Effect of Data Sampling When Using Random Forest on Imbalanced Bioinformatics Data	457
<i>David J. Dittman, Taghi M. Khoshgoftaar, and Amri Napolitano</i>	

IEEE International Workshop on Data Integration and Mining (DIM 2015)

Soft-Hard Clustering for Multiview Data	464
<i>Gaurav Tyagi, Nilesch Patel, and Ishwar Sethi</i>	
Trajectory Segmentation Using Oblique Envelopes	470
<i>Sudheer Chelluboina and Michael Hahsler</i>	
Efficacy of Season Prediction for Geo-locations Using Geo-tagged Images	476
<i>Shesha Sreenivasamurthy and Shayna Frank</i>	
Improving Pharmacological Research of HIV-1 Integrase Inhibition Using Differential Evolution - Binary Particle Swarm Optimization and Nonlinear Adaptive Boosting Random Forest Regression	485
<i>Richard Adrian Galvan, Ahmad Reza Hadaegh, and Matinehalsadat Kashani Moghaddam</i>	

IEEE International Workshop on Information Integration in Cyber Physical Systems (IICPS 2015)

A Low-Cost Haptic System for Wrist Rehabilitation	491
<i>Daniela D'Auria, Fabio Persia, and Bruno Siciliano</i>	
Web-Based Smart Grid Network Analytics Framework	496
<i>Stan Pietrowicz, Ben Falchuk, Aleksandar Kolarov, and Aditya Naidu</i>	

Session C2

Session C21: Information Security and Privacy II

Cyberdefense When Attackers Mimic Legitimate Users: A Bayesian Approach	502
<i>Atin Basuchoudhary, Mohamed Eltoweissy, Mohamed Azab, Laura Razzolini, and Shimaa Mohamed</i>	
Toward a Moving Target Defense for Web Applications	510
<i>Marthony Taguinod, Adam Doupé, Ziming Zhao, and Gail-Joon Ahn</i>	
Implementing and Managing Policy Rules in Attribute Based Access Control	518
<i>Vincent Hu, David F. Ferraiolo, D. Richard Kuhn, Raghu N. Kacker, and Yu Lei</i>	

Session C22: Ontologies & Integration

Towards the Cultural Knowledge Ontology	526
<i>Othusitse S.D. Phefo, Naledi Kefitiley, and Hlomani Hlomani</i>	
Computational Cost of Querying for Related Entities in Different Ontologies	534
<i>Chung Ming Cheung, Yinuo Zhang, Anand Panangadan, and Viktor K. Prasanna</i>	
Enhancing Efficiency of Web Search Engines through Ontology Learning from Unstructured Information Sources	542
<i>Eslam Amer</i>	

IEEE International Workshop on Empirical Methods for Recognizing Inference in Text (EMRITE 2015)

Ranking Online Customer Reviews with the SVR Model	550
<i>Hsien-You Hsieh and Shih-Hung Wu</i>	
Finding the Key Users in Facebook Fan Pages via a Clustering Approach	556
<i>Kuan-Cheng Lin, Shih-Hung Wu, Liang-Pu Chen, and Ping-Che Yang</i>	
Machine Learning for Imbalanced Datasets of Recognizing Inference in Text with Linguistic Phenomena	562
<i>Min-Yuh Day and Cheng-Chia Tsai</i>	

Neural Network-Based Vector Representation of Documents for Reader-Emotion Categorization	569
<i>Yu-Lun Hsieh, Shih-Hung Liu, Yung-Chun Chang, and Wen-Lian Hsu</i>	

IEEE International Workshop on Data Integration and Mining (DIM 2015)

Unconstrained Linear Spectral Mixture Models for Spatial Information Extraction: A Comparative Study	574
<i>Uttam Kumar, Cristina Milesi, S. Kumar Raja, Sangram Ganguly, and Ramakrishna R. Nemani</i>	
Link Analysis of Wikipedia Documents Using MapReduce	582
<i>Vasa Hardik, Vasudevan Anirudh, and Palanisamy Balaji</i>	

IEEE International Workshop on Information Integration in Cyber Physical Systems (IICPS 2015)

Cyber Physical System Integration and Configuration Guided by Satisfiability Modulo Theories	589
<i>Kishore Pochiraju and Sanjai Narain</i>	
Big Data Computing for Digital Forensics on Industrial Control Systems	593
<i>Julian Rrushi and Philip A. Nelson</i>	
State Estimation of a Distribution System Using WLS and EKF Techniques	609
<i>Faridoon Shabani, Masoumeh Seyedyazdi, Mohanmad Vaziri, Mahyar Zarghami, and Suresh Vadhva</i>	
Author Index	614