

## Table of Contents

Complexity Analysis in the Language of Information Technology Companies <i>Mary Luz Mouronte Lopez</i>	1
How to Find Important Users in a Web Community? Mining Similarity Graphs <i>Clemens Schefels</i>	10
Characterization of Network Traffic Data: A Data Preprocessing and Data Mining Application <i>Esra Kahya-Ozyirmidokuz, Ali Gezer, and Cebrail Ciflikli</i>	18
Structured Data and Source Code Analysis for Financial Fraud Investigations <i>Joe Sremack</i>	24
Integrity, or Lack Thereof, of the Electronic Record Systems of the Courts of the State of Israel <i>Joseph Zernik</i>	31
Network Visualization of Car Inspection Data using Graph Layout <i>Jaakko Talonen, Miki Sirola, and Mika Sulkava</i>	39
Trend Visualization on Twitter: What's Hot and What's Not? <i>Sandjai Bhulai, Peter Kampstra, Lidewij Kooiman, Ger Koole, Marijn Deurloo, and Bert Kok</i>	43
Travel Time Estimation Results with Supervised Non-parametric Machine Learning Algorithms <i>Ivana Cavar, Zvonko Kavran, and Ruder Michael Rapajic</i>	49
Automated Predictive Assessment from Unstructured Student Writing <i>Norma C. Ming and Vivienne Ming</i>	57
Building OLAP Data Analytics by Storing Path-Enumeration Keys into Sorted Sets of Key-Value Store Databases <i>Luis Loyola, Fernando Wong, and Daniel Pereira</i>	61
Ontology-Guided Data Acquisition and Analysis <i>Dominic Girardi, Klaus Arthofer, and Michael Giretzlehner</i>	71
Analysis of Streaming Service Quality Using Data Analytics of Network Parameters <i>Jie Zhang, Hwa-Jong Kim, and Doo-Heon Ahn</i>	76
Modeling Team-Compatibility Factors Using a Semi-Markov Decision Process: A Data-Driven Framework for Performance Analysis in Soccer <i>Ali Jarvandi, Shahram Sarkani, and Thomas Mazzuchi</i>	80

Design and Operation of the Electronic Record Systems of the US Courts are Linked to Failing Banking Regulation <i>Joseph Zernik</i>	83
The Open Data Interface (ODI) Framework for Public Utilization of Big Data <i>Hwa-Jong Kim, Seung-Teak Lee, and Yi-Chul Kang</i>	94
Evaluating Data Minability Through Compression - An Experimental Study <i>Dan Simovici, Dan Pletea, and Saaid Baraty</i>	97
A New Measure of Rule Importance Using Hellinger Divergence <i>Chang-Hwan Lee</i>	103
An Architecture for Semantically Enriched Data Stream Mining <i>Andreas Textor, Fabian Meyer, Marcus Thoss, Jan Schaefer, Reinhold Kroeger, and Michael Frey</i>	107