

# **2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI 2020)**

**Las Vegas, Nevada, USA  
11 – 13 August 2020**



**IEEE Catalog Number: CFP20IRI-POD  
ISBN: 978-1-7281-1055-4**

**Copyright © 2020 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP20IRI-POD
ISBN (Print-On-Demand):	978-1-7281-1055-4
ISBN (Online):	978-1-7281-1054-7

**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  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI) **IRI 2020**

## Table of Contents

Foreword	xiii
Message from the Program Co-Chairs	xiv
Organizing Committee	xvi
Program Committee	xvii

### Session A11: AI and Security

An Empirical Analysis on the Usability and Security of Passwords	1
<i>Kanwardeep Singh Walia (California State University, Sacramento), Shweta Shenoy (KLA Corporation), and Yuan Cheng (California State University, Sacramento)</i>	
Detection Methods of Slow Read DoS Using Full Packet Capture Data	9
<i>Clifford Kemp (Florida Atlantic University), Chad Calvert (Florida Atlantic University), and Taghi M. Khoshgoftaar (Florida Atlantic University)</i>	
RNN-VED for Reducing False Positive Alerts in Host-based Anomaly Detection Systems	17
<i>Lydia Bouzar-Benlabiod (Ecole nationale Supérieure d'Informatique), Stuart H. Rubin (Naval Information Warfare Center Pacific), Kahina Belaidi (ESI), and Nour ElHouda Haddar (ESI)</i>	
Addressing Imbalanced Data Problem with Generative Adversarial Network For Intrusion Detection	25
<i>Ibrahim Yilmaz (Tennessee Tech University)</i>	

### Session A12: Computer Security and Privacy I

Combating Hard or Soft Disasters with Privacy-Preserving Federated Mobile Buses-and-Drones Based Networks	31
<i>Bo Ma (Auckland University of Technology), Jinsong Wu (Universidad de Chile, Guilin University of Electronic Technology), William Liu (Auckland University of Technology), Luca Chiaraviglio (Auckland University of Technology), and Xing Ming (Auckland University of Technology)</i>	

Relating the Empirical Foundations of Attack Generation and Vulnerability Discovery .37.....	
	<i>Tyler Westland (University of Cincinnati), Nan Niu (University of Cincinnati), Rashmi Jha (University of Cincinnati), David Kapp (AFRL/Rywa), and Temesguen Kebede (AFRL/Rywa)</i>
Privacy-Preserving and Efficient Sharing of Drone Videos in Public Safety Scenarios using Proxy Re-Encryption .45.....	
	<i>Khaled Rabieh (Metropolitan State University), Suat Mercan (Florida International University), Kemal Akkaya (Florida International University), Vashish Baboolal (Florida International University), and Ramazan S. Aygun (Kennesaw State University)</i>
Relevance of Grapheme's Shape Complexity in Writer Verification Task .53.....	
	<i>Ameur Bensefia (Higher Colleges of Technology) and Chawki Djeddi (Larbi Tebessi University)</i>

## Session A21: Computer Vision

Multimodal Information Integration for Indoor Navigation Using a Smartphone .59.....	
	<i>Yaohua Chang (City University of New York), Jin Chen (City University of New York), Tyler Franklin (City University of New York), Lei Zhang (Borough of Manhattan Community College – CUNY), Arber Ruci (New York City Regional Innovation Node – CUNY), Hao Tang (Borough of Manhattan Community College – CUNY), and Zhigang Zhu (The City University of New York)</i>
Automated Filtering of Eye Gaze Metrics from Dynamic Areas of Interest .67.....	
	<i>Gavindya Jayawardena (Old Dominion University) and Sampath Jayarathna (Old Dominion University)</i>
Dynamic Image for Micro-Expression Recognition on Region-Based Framework .75.....	
	<i>Trang Thanh Quynh Le (University of St. Thomas), Thuong Khanh Tran (University of Oulu), and Manjeet Rege (University of St. Thomas)</i>
Building Damage Evaluation from Satellite Imagery using Deep Learning .82.....	
	<i>Fei Zhao (University of Alabama at Birmingham) and Chengcui Zhang (University of Alabama at Birmingham)</i>

## Session A22: Computer Security and Privacy II

Distributed Differentially Private Mutual Information Ranking and Its Applications .90.....	
	<i>Ankit Srivastava (Microsoft), Samira Pouyanfar (Microsoft), Joshua Allen (Microsoft), Ken Johnston (Microsoft), and Qida Ma (Microsoft)</i>
Medicare Fraud Detection using CatBoost .97.....	
	<i>John Hancock (Florida Atlantic University) and Taghi M. Khoshgoftaar (Florida Atlantic University)</i>
A Distribution-based Regression for Real-Time COVID-19 Cases Detection from Chest X-Ray and CT Images .104.....	
	<i>Nuha Zamzami (University of Jeddah), Pantea Koochemeshkian (Concordia University), and Nizar Bouguila (Concordia University)</i>

## Session A31: Machine Learning and Data Mining I

Background Subtraction with a Hierarchical Pitman-Yor Process Mixture Model of Generalized Gaussian Distributions .112.....	112
<i>Srikanth Amudala (Concordia University), Samr Ali (Concordia University), and Nizar Bouguila (Concordia University)</i>	
Fully Bayesian Learning of Multivariate Beta Mixture Models .120.....	120
<i>Mahsa Amirkhani (Concordia University), Narges Manouchehri (Concordia University), and Nizar Bouguila (Concordia University)</i>	
Post-Click Behaviors Enhanced Recommendation System .128.....	128
<i>Zhenhua Liang (New York Institute of Technology), Siqi Huang (University of North Carolina at Charlotte), Xueqing Huang (New York Institute of Technology), Rui Cao (New Jersey Institute of Technology), and Weize Yu (Shandong University)</i>	
The Democratization of Machine Learning Features .136.....	136
<i>Jayesh Patel (Rockstar Games)</i>	
Identifying Feature Pattern for Weighted Imbalance Data: A Feature Selection Study for Thoracolumbar Spine Fractures in Crash Injury Research .142.....	142
<i>Paromita S. Nitu (Marquette University), Praveen Madiraju (Marquette University), and Frank A. Pintar (Marquette University &amp; Medical College of Wisconsin)</i>	

## Session A32: Data & Knowledge Representation and Management

AD4ML: Axiomatic Design to Specify Machine Learning Solutions for Manufacturing .148.....	148
<i>Alejandro Gabriel Villanueva Zacarias (University of Stuttgart), Rachaa Ghabri (University of Stuttgart), and Peter Reimann (University of Stuttgart)</i>	
Data Driven Relational Constraint Programming .156.....	156
<i>Michael Valdron (Ontario Tech University) and Ken Q. Pu (Ontario Tech University)</i>	
Quality not Quantity! A Qualitative Evaluation and Proposal for Understanding the Depth of Audience "Knowledge" Post Data Extraction .164.....	164
<i>Kimberley Hemmings-Jarrett (Drexel University), Terryann Barnett (NYC Department of Education), Julian Jarrett (Lutron Electronics), M. Brian Blake (George Washington University), and Denise Agosto (Drexel University)</i>	
Approximate Matching of Spatiotemporal RDF Data by Path .172.....	172
<i>Jiajia Lu (Northeastern University), Xiaofeng Di (Northeastern University), and Luyi Bai (Northeastern University)</i>	

## Session B11: AI and Medical Informatics

- A Novel AI-Enabled Framework to Diagnose Coronavirus COVID-19 using Smartphone Embedded Sensors: Design Study .180.....  
*Halgurd S. Maghded (Koya University), Kayhan Zrar Ghafoor (Salahaddin University-Erbil), Ali Safaa Sadiq (University of Wolverhampton), Kevin Curran (Ulster University), Danda B. Rawat (Howard University), and Khaled Rabie (Manchester Metropolitan University)*
- Attention-Guided Generative Adversarial Network to Address Atypical Anatomy in Synthetic CT Generation .188.....  
*Hajar Emami (Wayne State University), Ming Dong (Wayne State University), and Carri K. Glide-Hurst (Henry Ford Health System)*
- Empowering Team Science Across the Translational Spectrum with the UAB Biomedical Research Infrastructure Technology Enhancement (U-BRITE) .194.....  
*James J. Cimino (University of Alabama at Birmingham), Wayne H. Liang (University of Alabama at Birmingham), Jelai Wang (University of Alabama at Birmingham), Dongmei Sun (University of Alabama at Birmingham), John D. Osborne (University of Alabama at Birmingham), Amy Y. Wang (University of Alabama at Birmingham), S. Louis Bridges Jr. (University of Alabama at Birmingham), Matt Wyatt (University of Alabama at Birmingham), and Jake Y. Chen (University of Alabama at Birmingham)*
- Predicting PTSD Severity in Veterans from Self-Reports for Early Intervention: A Machine Learning Approach .201.....  
*Priyanka Annatureddy (Marquette University), Md Fitrat Hossain (Marquette University), Thomas Kissane (Marquette University), Wylie Frydrychowicz (Marquette University), Paromita Nitu (Marquette University), Joseph Coelho (Marquette University), Nadiyah Johnson (Marquette University), Praveen Madiraju (Marquette University), Zeno Franco (Medical College of Wisconsin), Katinka Hooyer (Medical College of Wisconsin), Niharika Jain (Marquette University), Mark Flower (Mental Health America), and Sheikh Ahamed (Marquette University)*

## Session B12: Predictive Modeling and Data Analysis

- Forecasting Atmospheric Visibility using Auto Regressive Recurrent Neural Network .209.....  
*Jahnavi Jonnalagadda (George Mason University) and Mahdi Hashemi (George Mason University)*
- Global Land Temperature Forecasting using Long Short-Term Memory Network .216.....  
*Prashanti Maktala (George Mason University) and Mahdi Hashemi (George Mason University)*
- Semantic Embeddings for Medical Providers and Fraud Detection .224.....  
*Justin M. Johnson (Florida Atlantic University) and Taghi M. Khoshgoftaar (Florida Atlantic University)*
- Studying the Impact of Streetlights on Street Crime Rate using Geo-Statistics .231.....  
*Srikanth Vadlamani (George Mason University) and Mahdi Hashemi (George Mason University)*

## Session B31: Machine Learning and Data Mining II

- Development of Sentiment Lexicon in Bengali Utilizing Corpus and Cross-Lingual Resources .237.  
*Salim Sazzed (Old Dominion University)*
- Multi-Label Multi-Task Learning with Dynamic Task Weight Balancing .245.....  
*Tianyi Wang (Florida International University) and Shu-Ching Chen (Florida International University)*
- Semantic Data Understanding with Character Level Learning .253.....  
*Michael J. Mior (Rochester Institute of Technology) and Ken Q. Pu (Ontario Tech University)*
- Natural Language-based Integration of Online Review Datasets for Identification of Sex Trafficking Businesses .259.....  
*Maria Diaz (California State University, Fullerton) and Anand Panangadan (California State University, Fullerton)*
- Using a Deep Learning Model, Content Features, and Author Metadata to Recommend Research Papers .265.....  
*Si-Hong Lam (Brigham Young University), Eric Brewer (Brigham Young University), and Yiu-Kai Ng (Brigham Young University)*

## Session B32: Database and Graph Models

- An Approach for Schema Extraction of NoSQL Graph Databases .271.....  
*Angelo Augusto Frozza (Instituto Federal Catarinense), Salomão Rodrigues Jacinto (Universidade Federal de Santa Catarina), and Ronaldo dos Santos Mello (Universidade Federal de Santa Catarina)*
- KGdiff: Tracking the Evolution of Knowledge Graphs .279.....  
*Abbas Keshavarzi (University of Georgia) and Krys J. Kochut (University of Georgia)*
- A Water Quality Research Platform for the Near-Real-Time Buoy Sensor Data .287.....  
*Maria Presa-Reyes (Florida International University), Biayna Bogosian (Florida International University), Bradley Schonhoff (Florida International University), Christopher Jerauld (Florida International University), Christian Moreyra (Florida International University), Piero Gardinali (Florida International University), and Shu-Ching Chen (Florida International University)*
- Towards an Accountability Suggestion Map for Supporting Information Systems Management Based on Systems Thinking .295.....  
*Felipe Cordeiro (UNIRIO), Aline Pires Vieira de Vasconcelos (IFF), Rodrigo Pereira dos Santos (UNIRIO), and Patricia Lago (Vrije Universiteit Amsterdam)*
- Wrapping a NoSQL Datastore for Stream Analytics .301.....  
*Khalid Mahmood (Uppsala University), Kjell Orsborn (Uppsala University), and Tore Risch (Uppsala University)*

## Session C11: Heuristic Acquisition and Reasoning

- An Adaptive and Dynamic Biosensor Epidemic Model for COVID-19 .306.....  
*Salvador V. Balkus (University of Massachusetts Dartmouth), Joshua Rumbut (University of Massachusetts Dartmouth, University of Massachusetts Medical School), Honggang Wang (University of Massachusetts Dartmouth), and Hua Fang (University of Massachusetts Dartmouth, University of Massachusetts Medical School)*
- Artificial Intelligence and Data Science Governance: Roles and Responsibilities at the C-Level and the Board .314.....  
*Bhavani Thuraisingham (The University of Texas at Dallas)*
- An I/O Request Packet (IRP) Driven Effective Ransomware Detection Scheme using Artificial Neural Network .319.....  
*Md. Ahsan Ayub (Tennessee Tech University), Andrea Continella (University of Twente), and Ambareen Siraj (Tennessee Tech University)*
- A New Emulation Platform for Real-Time Machine Learning in Substance Use Data Streams .325...  
*Stefan A. Bruendl (University of Massachusetts Dartmouth), Hua Fang (University of Massachusetts Dartmouth), Hieu Ngo (University of Massachusetts Dartmouth), Edward W. Boyer (Harvard Medical School), and Honggang Wang (University of Massachusetts Dartmouth)*

## Session C12: Data Integration and Mining (DIM) Workshop

- Towards Agile Integration: Specification-based Data Alignment .333.....  
*Chris Giossi (Portland State University), David Maier (Portland State University), Kristin Tufte (Portland State University), Elliot Gall (Portland State University), and Melissa Barnes (Portland State University)*
- Fairness in Data Wrangling .341.....  
*Lacramioara Mazilu (University of Manchester), Norman W. Paton (University of Manchester), Nikolaos Konstantinou (University of Manchester), and Alvaro A.A. Fernandes (University of Manchester)*
- Latent Feature Modelling for Recommender Systems .349.....  
*Abdullah Alhejaili (Loughborough University) and Shaheen Fatima (Loughborough University)*
- Mining Frequent Differences in File Collections .357.....  
*Sudarshan S. Chawathe (University of Maine)*

## Session C21: Artificial Intelligence for HealthCare (AIHC) Workshop

- An Intelligent Baby Monitoring System Based on Raspberry PI, IoT Sensors and Convolutional Neural Network .365.....  
*Rabea Cheggou (Laboratoire des Technologies Innovantes, Ecole Nationale Supérieure de Technologie), Siham Si hadj mohand (Laboratoire des Technologies Innovantes, Ecole Nationale Supérieure de Technologie), Oussama Annad (Laboratoire des Technologies Innovantes, Ecole Nationale Supérieure de Technologie), and El hadi Khoumeri (Laboratoire des Technologies Innovantes, Ecole Nationale Supérieure de Technologie)*



Multi-Class Cardiovascular Diseases Diagnosis from Electrocardiogram Signals using 1-D Convolution Neural Network .372.....  
*Mehdi Fasihi (IAU), Mohammad H. Nadimi-Shahraki (Islamic Azad University), and Ali Jannesari (Iowa State University IA, USA)*

Toward Data-Driven Assessment of Caregiver's Burden for Persons with Dementia using Machine Learning Models .379.....  
*Hilda Goins (North Carolina A&T State University), SeyyedPooya HekmatiAthar (North Carolina A&T State University), Grace Byfield (North Carolina A&T State University), Raymond Samuel (North Carolina A&T State University), and Mohd Anwar (North Carolina A&T State University)*

Discovering Drug-Drug and Drug-Disease Interactions Inducing Acute Kidney Injury Using Deep Rule Forests .385.....  
*Bowen Kuo (National Sun Yat-sen University), Yihuang Kang (National Sun Yat-sen University), Pinghsung Wu (Kaohsiung Medical University), Sheng-Tai Huang (National Sun Yat-sen University), and Yajie Huang (National Sun Yat-sen University)*

## Session C22: Poster and Demo Session

DataOps for Societal Intelligence: a Data Pipeline for Labor Market Skills Extraction and Matching .391.....  
*Damian A. Tamburri (Technical University Eindhoven, Jheronimus Academy of Data Science), Willem-Jan Van Den Heuvel (Tilburg University, Jheronimus Academy of Data Sciences), and Martin Garriga (Tilburg University, Jheronimus Academy of Data Sciences)*

NLP Relational Queries and Its Application .395.....  
*Andrei Stoica (Ontario Tech University), Ken Q. Pu (Ontario Tech University), and Heidar Davoudi (Ontario Tech University)*

Using Deep Learning To Assign Rheumatoid Arthritis Scores .399.....  
*Son Do Hai Dang (University of Alabama at Birmingham) and Leigh Allison (University of Alabama at Birmingham)*

Development of Natural Language Processing Algorithm for Dental Charting .403.....  
*Yifan Zhang (University of Alabama at Birmingham), Brandon Bogard (Berry College), and Chengcui Zhang (University of Alabama at Birmingham)*

Topic Diffusion Discovery Based on Deep Non-Negative Autoencoder .405.....  
*Sheng-Tai Huang (National Sun Yat-sen University), Yihuang Kang (National Sun Yat-sen University), Shao-Min Hung (National Sun Yat-sen University), Bowen Kuo (National Sun Yat-sen University), and I-Ling Cheng (National Chung Hsing University)*

## Session C31: Empirical Methods for Recognizing Inference in Text (EM-RITE) Workshop

- Modeling of Clinical Mammography Recognition 409.....  
*Kuo-Chung Chu (National Taipei University of Nursing and Health Sciences), Po-Yao Tsai (National Taipei University of Nursing and Health Sciences), Tien-Yu Chang (Cheng Hsin General Hospital), and Yu-Shu Wu (Cheng Hsin General Hospital)*
- Cross-Domain Helpfulness Prediction of Online Consumer Reviews by Deep Learning Model 412  
*Shih-Hung Wu (Chaoyang University of Technology) and Yi-Kun Chen (Chaoyang University of Technology)*
- A Study of Deep Learning for Factoid Question Answering System 419.....  
*Min-Yuh Day (National Taipei University) and Yu-Ling Kuo (Tamkang University)*
- Research on Online Impulsive Buying and Post-Purchase Dissonance 425.....  
*Wen-Kuo Chen (Chaoyang University of Technology), Yen-Ling Lin (Chaoyang University of Technology), Hua-Sheng Pan (Chaoyang University of Technology), and Cheng-Kun Chen (Chaoyang University of Technology)*
- Building the Body Image Conceptual Framework Based on the Theory of Planned Behavior (TPB) 430  
*Wen-Kuo Chen (Chaoyang University of Technology), Ying-Hsun Hung (Hwa Hsia University of Technology), and Jun-Yu Zhong (Chaoyang University of Technology)*

## Session C32: Novel Applications

- Ultra Wideband Indoor Positioning System Based on Artificial Intelligence Techniques 438.....  
*Long Cheng (ABB Inc), Zhaoqi Wu (University of Illinois at Urbana-Champaign), Bo Lai (University of California Irvine), Qiang Yang (North Carolina State University), Anguo Zhao (University of California Irvine), and Yuanting Wang (Rensselaer Polytechnic Institute)*
- BusinessDetect: An Advanced Business Information Mining Application for Intelligent Marketing 445.....  
*Ye Qiu (Peking University), Xiaolong Gong (Peking University), and Zhiyi Ma (Peking University)*
- A Comparison of Machine Learning Algorithms Applied to American Legislature Polarization 451  
*Gabriel Mersy (University of Minnesota), Vincent Santore (The Burr Project), Isaac Rand (University of Chicago), Corrine Kleinman (Georgetown University), Grant Wilson (Tulane University), Jason Bonsall (University of Nevada), and Tyler Edwards (Temple University)*
- The EGRBAC Model for Smart Home IoT 457.....  
*Safwa Ameer (The University of Texas at San Antonio), James Benson (The University of Texas at San Antonio), and Ravi Sandhu (The University of Texas at San Antonio)*
- Author Index 463** .....