



IEEE BigData 2020 Program Schedule Virtual Event

Conference time zone: USA Eastern Standard Time

A few days before the event, every registered participant will get an email from the company to provide the login information to access the virtual conference platform.

Sponsored by



2020 IEEE International Conference on Big Data

IEEE Big Data 2020 Program Schedule	2
Keynote Lectures	8
Conference Paper Presentations	10
Industry and Government Paper Presentations	16
Tutorials	19
Workshops.....	22
Special Symposiums	24
Special Sessions	25
BigData Cup Challenges	53
Posters.....	61

IEEE Big Data 2020 Program Schedule

Virtual Event

December 10 - December 13, 2020

Keynote Lecture: **60 minutes** (about 45 minutes for talk and 15 minutes for Q and A)
Main conference regular paper: **25 minutes** (about 20 minutes for talk and 5 minutes for Q and A)
Main conference short paper: **15 minutes** (about 11 minutes for talk and 4 minutes for Q and A)

All conference activities take place virtually and are based on USA Eastern Standard time.

Time	Sessions/Workshops	Session Chair
9am-6pm	BigData Cup Challenges: Global Road Damage Detection Challenge 2020	Hiroya Maeda
	Information Granulation in Data Science and Scalable Computing	Shusaku Tsumoto, Dominik Slezak, Tzung-Pei Hong Shyue-Liang Wang Weiping Ding
	The Fourth Annual Workshop on Applications of Artificial Intelligence in the Legal Industry	Haozhen Zhao
	The 4th International Workshop on Big Data Analytic for Cyber Crime Investigation and Prevention	Andrii Shalaginov
	The 4th Workshop on Benchmarking, Performance Tuning and Optimization for Big Data Applications (BPOD)	Zhiyuan Chen
Time	Sessions/Workshops	Session Chair
Morning (9am-1:30pm)	Special Session 6: Explainable Artificial Intelligence in Safety Critical Systems	Yixin Chen, Lixin Fan, Qiang Yang
9am-1:30pm	Second Workshop on Big Data Predictive Maintenance Using Artificial Intelligence	Rituparna Datta
9am-1:30pm	9th Workshop on Scalable Cloud Data Management	Benjamin Wollmer
	6th International Workshop to Improve Big Data Science Project Team Processes	Jeffrey Saltz
	Workshop on IoT based Big Data Architectures and Applications	Hasan Ali Khattak
	International Workshop on Fair and Interpretable Learning Algorithms (FILA 2020)	Mayank Singh
	2nd IEEE Workshop on Machine Learning for Big Data Analytics in Remote Sensing	Maryam Rahnemoonfar
	Big Data Analytic Technology for Bioinformatics and Health Informatics (KDDDBHI)	Xin Deng
	Advances in High Dimensional (AdHD) Big Data	Sotiris Tasoulis
	Applications of Big Data Technology in the Transport Industry	John Easton

	3rd Workshop on Big Data for CyberSecurity (BigCyber-2020)	Sudip Mittal
	Discriminative Pattern Mining Workshop (DPM 2020)And 4th International Workshop on Big Data Analytics for Cyber Intelligence and Defense (BDA4CID 2020)	<i>Ryan Benton, Tom Johnsten, Suresh Choubey, Stephen McGough and Amir Atapour Abarghouei</i>
9:00-2:30pm	The 5th Workshop on Real-time Stream Analytics, Stream Mining, CER/CEP & Stream Data Management in Big Data	Sabri Skhiri
9:00-2:30pm	International Workshop on Big Data Reduction	Dingwen
9:00-2:30pm	The 2nd International Workshop on Big Data Tools, Methods, and Use Cases for Innovative Scientific Discovery (BTSD)	Sangkeun (Matt)
9-2pm	The 5th IEEE International Workshop on Big Spatial Data (BSD 2020)	Farnoush Banaei-Kashani
9-2pm	The 4th IEEE International Workshop on Big Data for Financial News and Data	Quanzhi Li
12~6pm	The 4th Workshop on Graph Techniques for Adversarial Activity Analytics	Jiejun Xu
Afternoon (1:30-6:00pm)	Workshop on Smart Farming, Precision Agriculture, and Supply Chain (SmartFarm-2020)	Sudip Mittal
1:30-6:00pm	3rd Workshop on Energy-Efficient Machine Learning and Big Data Analytics	Mohammed Alawad
1:30-6:00pm	IEEE Big Data Workshop on Data Science for Smart and Connected Communities	George Mohler
1:30-6:00pm	The 3rd International Workshop on Big Media Dataset Construction, Management and Applications	Cheng Jin
1:30-6:00pm	Workshop on Deviant Activities on Social Media	Muhammad Nihal Hussain
1-6pm	The fourth workshop on Human-in-the-loop Methods and Future of Work in BigData	Senjuti Basu Roy, Alex Quinn, and Atsuyuki Morihsima
10:45-6:00PM	International Workshop on Data Analytics for Smart Health (DASH 2020)	Harishchandra.Dubey@microsoft.com
Time	Sessions/Workshops	Session Chair
9-11am	Tutorial 5: A Tutorial on Topological Data Analysis in Text Mining	
9-11am	Tutorial 9: Big Sequence Management	

9:00-11:10AM	Poster Session 1 (Poster P201-P220)	Anuj Karpatne, Jianwu Wang
11:20-1:30pm	Poster Session 2 (Poster P221-P238)	Anuj Karpatne, Jianwu Wang
1:30-3:30pm	Tutorial 7: Big data stream mining	Bartosz Krawcz Alberto Cano

Day 2: Friday, December 11, 2020		
8: 30 – 9:00 am Location:	Opening and Welcome Conf Chairs, PC Chairs, I&G Chair	
9:00 – 10:00 am Location:	Keynote: Prof. Jeannette M. Wing Session Chair: Christopher.M. Jermaine	
10: 00 - 10: 15 am Location:	Coffee Break	
Time	Sessions/Workshops	Session Chair
9am-6pm	3rd Special Session on HealthCare Data	Sultan Turhan, Ozgun Pinarer
	6th Special Session on Intelligent Data Mining	Uraz Yavanoglu
9am-6pm	Computational Archival Science: digital records in the age of big data	Mark Hedges
9am-6pm	2nd Special Session on Machine Learning on Big Data	Alfredo Cuzzocrea
9am-6pm	2020 Workshop on Data Science in Medicine and Healthcare	Xiong Liu
9:00am-1:30pm	8th International Workshop on Distributed Storage and Blockchain Technologies for Big Data	Hui Li
9:00am-1:30pm	International Workshop on Big Data Analytics for Cyber Threat Hunting (CyberHunt 2020)	Vasileios Mavroeidis
9:00am-1:30pm	The 7th Workshop on Performance Engineering with Advances in Software and Hardware for Big Data Sciences (PEASH)	Hui Zhang
9:00am-1:30pm	BIG FOOD AND NUTRITION DATA MANAGEMENT AND ANALYSIS	Tome Eftimov
9:00am-1:30pm	Discriminative Pattern Mining Workshop	Ryan Benton
10:15 – 12:20 pm	S1: Systems for High Performance ML and Numerical Computing (3 full + 3 short)	Yi Zhou
	S2: Workflows and Monitoring (2 full + 5 short)	ChengXiang Zhai
	S3: Big Human-Derived Data (3 full + 3 short)	Xintao Wu
	S4: Big Transportation Data (3 full + 3 short)	Yannis Velegrakis
	I&G 1: Big Data: Machine Learning Applications I	Zhongfang Zhuang
	Tutorial 8: Learning from complex medical data"	Panagiotis Papapetrou Jaakko Hollmen

	Myra Spiliopoulou	
10:15 – 12:20 pm	Poster Session 3 (Poster P239-P255)	Anuj Karpatne, Jianwu Wang, jianwu
12: 20 - 1: 30 pm Location:	Lunch break	
Time	Sessions/Workshops	Session Chair
1:30 – 3:35 pm	S5: Computational Models for Big Data (5 full)	Xinghui Zhao
	S6: Data Analytics Engines (3 full + 3 short)	Jia Zou
	S7: Big Data Systems (3 full + 3 short)	Sang-Woo Jun
	S8: Big Biological and Physical Data (2 full + 5 short)	Xiaohua Tony HU
	I&G 2: Big Data: Machine Learning Applications II	Weijia Xu
	Tutorial 2: Big Data System Benchmarking: State of the Art, Current Practices, and Open Challenges	Chao Zhang, Jiaheng Lu
3: 35 - 3: 55 pm Location:	Coffee Break	
Time	Sessions/Workshops	Session Chair
3:55 - 6: 00 pm	S9: Big Human-Derived Data (3 full + 3 short)	ChengXiang Zhai
	S10: Big Transportation and Spatial Data (3 full + 3 short)	Aris Gkoulalas-Divanis
	S11: Medical Data (3 full + 3 short)	Thirimachos Bourlai
	S12: Theoretical Models for Big Data (3 full + 3 short)	Carlos Ordonez
	I&G 3: Big Data: Feature and Anomaly Detection I	Nancy Grady
	Tutorial 6: Data Sources, Tools, and Techniques for Big Data-driven Machine Learning in Heliophysics	Dustin Kempton, Azim Ahmadzadeh

Day 3: Saturday, December 12, 2020		
9:00 - 10: 00 am Location:	Keynote: Prof. Christos Faloutsos Session Chair: Xintao Wu	
10: 00 - 10: 15 am Location:	Coffee Break	
9am-1pm	2nd Special Session on Machine Learning on Big Data	Alfredo Cuzzocrea
9-1:30pm	Innovation Workshop on Transforming Big Data into Actionable Knowledge (BiDAW)	Mahdi Bohlouli
9-1:30pm	User Understanding for Big Data, 2020	Claire Ding
9-1:30pm	Seventh International Workshop on High Performance Big Graph Data Management, Analysis, and Mining (BigGraphs 2020)	alhasan@iupui.edu
9am-6pm	Third International Workshop of Internet of Things Data Analytics (IoTDA)	Eyhab Al-Masri
Time	Sessions/Workshops	Session Chair
10: 15 - 12: 20 pm	S13: Health and Human Motion Data (3 full + 3 short)	xiaoqian jiang
	S14: Theoretical and Computational Models for Big Data (2 full + 5 short)	Siamak Yousefi
	S15: Computational Models for Big Data (2 full + 5 short)	Xiaohua Tony Hu
	S16: Efficient and Scalable Data Management and Mining (3 full + 3 short)	Samuel A. Ajila
	I&G 4: Big Data: Feature and Anomaly Detection II	Wang Zhou
	Tutorial 1: Big Data Applications for High-Impact / Emerging Topics: Case Studies on Molecular Biology and Its Applications to COVID-19	Jiecong Lin Shixiong Zhang
	Poster Session 4 (Poster P256-P273)	Anuj Karpatne, Jianwu Wang
12: 20 - 1: 30 pm Location:	Lunch break	
Time	Sessions/Workshops	Session Chair
1: 30 - 3: 35 pm	S17: Multi-structured Data - Big Variety Data (3 full + 3 short)	ChengXiang Zhai
	S18: Data Streams and Time Series (3 full + 3 short)	Carson Leung
	S19: Web Search and Mining (3 full + 3 short)	Tommy Dang
	S20: Recommendation and Graph Mining (5 full)	xiaoqian jiang
	I&G 5: Big Data ML Platforms and Systems I	Siyuan Lu
	Tutorial 3: Systems and Algorithms for Massively Parallel Graph Mining	Da Yan Guimu Guo
3: 35 - 3: 55 pm	Coffee Break	
Time	Sessions/Workshops	Session Chair
3: 55 - 6: 00 pm	S21: Data Preprocessing and Integration (5 full)	Anna Squicciarini
	S22: Big Variety Data (3 full + 3 short)	Xintao Wu

	S23: Threat Detection and Web Search (2 full + 5 short)	Thirimachos Bourlai
	S24: Privacy Preserving Big Data Collection/Analytics (4 full + 1 short)	xiaoqian jiang
	I&G 6: Big Data ML Platforms and Systems II	Dhaval Patel

Day 4: Sunday, December 13, 2020		
9: 00 - 10: 00 am Location:	Keynote: Dr. Divesh Srivastava Session Chair: Li Xiong	
10:00 – 10:15 am	Coffee Break	
Time	Sessions/Workshops	Session Chair
10:15 - 12: 20 am	S25: Recommender Systems and Graphs (3 full + 3 short)	Xintao Wu
	S26: Graph Mining and Embedding (3 full + 3 short)	Edoardo Serra
	S27: Temporal and StreamData Mining (2 full + 5 short)	Jie Ding
	S28: Representation Learning (2 full + 5 short)	Alfredo Cuzzocrea
	I&G 7: Data Quality and Privacy	Olivera Kotevska
10:15 - 12: 20 am	Tutorial 4: Semantic Exploration of Big Data	Maria Krommyda Verena Kantere
10:15 - 12: 20 am	Poster Session 5 (Poster P274-290)	Anuj Karpatne, Jianwu Wang
9:00-1:30pm	The 6th International Workshop on Solar & Stellar Astronomy Big Data (SABiD)	Rafal Angryk
9:00-1:30pm	The 3rd International Workshop on “Big Data Engineering and Analytics in Cyber-Physical Systems” (BigEACPS)	Akbar Namin
9:00-1:30pm	7th International Workshop on Privacy and Security of Big Data (PSBD 2020)	Alfredo Cuzzocrea
12: 20 - 12:45 pm Location:	Closing Ceremony	

Keynote Lectures

Keynote 1: Data for Good: Ensuring the Responsible Use of Data to Benefit Society

Speaker:

Jeannette M. Wing, Avaneessians Director of the Data Science Institute, Professor of Computer Science, Columbia University, USA

Abstract:

Every field has data. We use data to discover new knowledge, to interpret the world, to make decisions, and even to predict the future. The recent convergence of big data, cloud computing, and novel machine learning algorithms and statistical methods is causing an explosive interest in data science and its applicability to all fields. This convergence has already enabled the automation of some tasks that better human performance. The novel capabilities we derive from data science will drive our cars, treat disease, and keep us safe. At the same time, such capabilities risk leading to biased, inappropriate, or unintended action. The design of data science solutions requires both excellence in the fundamentals of the field and expertise to develop applications which meet human challenges without creating even greater risk. The Data Science Institute at Columbia University promotes “Data for Good”: using data to address societal challenges and bringing humanistic perspectives as—not after—new science and technology is invented. Started in 2012, the Institute is now a university-level institute representing over 350 affiliated faculty from 16 different schools and institutes across campus. Data science literally touches every corner of the university. In this talk, I will present the mission of the Institute and highlights of our educational and research activities—all with the aim of ensuring the responsible use of data to benefit society.

Short Bio:

Prof. Jeannette M. Wing is Avaneessians Director of the Data Science Institute and Professor of Computer Science at Columbia University. From 2013 to 2017, she was a Corporate Vice President of Microsoft Research. She is Adjunct Professor of Computer Science at Carnegie Mellon where she twice served as the Head of the Computer Science Department and had been on the faculty since 1985. From 2007-2010 she was the Assistant Director of the Computer and Information Science and Engineering Directorate at the National Science Foundation. She received her S.B., S.M., and Ph.D. degrees in Computer Science, all from the Massachusetts Institute of Technology. Professor Wing's general research interests are in the areas of trustworthy computing, specification and verification, concurrent and distributed systems, programming languages, and software engineering. Her current interests are in the foundations of security and privacy, with a new focus on trustworthy AI. She was or is on the editorial board of twelve journals, including the Journal of the ACM and Communications of the ACM.

Professor Wing is known for her work on linearizability, behavioral subtyping, attack graphs, and privacy-compliance checkers. Her 2006 seminal essay, titled Computational Thinking is credited with helping to establish the centrality of computer science to problem-solving in fields where previously it had not been embraced. She is currently a member of: the National Library of Medicine Blue Ribbon Panel; the Science, Engineering, and Technology Advisory Committee for the American Academy for Arts and Sciences; the Board of Trustees for the Institute of Pure and Applied Mathematics; the Advisory Board for the Association for Women in Mathematics; and the Alibaba DAMO Technical Advisory Board. She has been chair and/or a member of many other academic, government, and industry advisory boards. She received the CRA Distinguished Service Award in 2011 and the ACM Distinguished Service Award in 2014. She is a Fellow of the American Academy of Arts and Sciences Association for the Advancement of Science, the Association for Computing Machinery (ACM), and the Institute of Electrical and Electronic Engineers (IEEE).

Keynote 2: Anomaly Detection in Large Graphs

Speaker:

Christos Faloutsos, Professor, Carnegie Mellon University, USA

Abstract:

Given a large graph, like who-calls-whom, or who-likes-whom, what behavior is normal and what should be surprising, possibly due to fraudulent activity? How do graphs evolve over time? We focus on these topics: (a) anomaly detection in large static graphs and (b) patterns and anomalies in large time-evolving graphs. For the first, we present a list of static and temporal laws, including advances patterns like 'eigenspokes'; we show how to use them to spot suspicious activities, in on-line buyer-and-seller settings, in FaceBook, in twitter-like networks. For the second, we show how to handle time-evolving graphs as tensors, as well as some surprising discoveries such settings.

Short Bio:

Christos Faloutsos is a Professor at Carnegie Mellon University and an Amazon Scholar. He received the Fredkin Professorship in Artificial Intelligence (2020); the Presidential Young Investigator Award by the National Science Foundation (1989), the Research Contributions Award in ICDM 2006, the SIGKDD Innovations Award (2010), the PAKDD Distinguished Contributions Award (2018), 28 “best paper” awards (including 7 “test of time” awards), and four teaching awards. Eight of his advisees or co-advisees have attracted KDD or SCS dissertation awards. He is an ACM Fellow, he has served as a member of the executive committee of SIGKDD; he has published over 400 refereed articles, 17 book chapters and three monographs. He holds 8 patents (and several more are pending), and he has given over 50 tutorials and over 25 invited distinguished lectures.

His research interests include large-scale data mining with emphasis on graphs and time sequences; anomaly detection, tensors, and fractals.

Keynote 3: Towards High-Quality Big Data: A Focus on Time

Speaker:

Divesh Srivastava, Head of Database Research, Data Science and AI Research, AT&T Chief Data Office, USA*Abstract:*

Data are being generated, collected, and analyzed today at an unprecedented scale, and data-driven decision making is sweeping through all aspects of society. As the use of big data has grown, so too have concerns that poor-quality data, prevalent in large data sets, can have serious adverse consequences on data-driven decision making. Responsible data science thus requires a recognition of the importance of veracity, the fourth “V” of big data. In this talk, we first present a vision of high-quality big data and highlight the substantial challenges that the first three V’s, volume, velocity, and variety, bring to dealing with veracity in big data. We then present the FIT Family of adaptive, data-driven statistical tools that we have designed, developed, and deployed at AT&T for continuous data quality monitoring of a large and diverse collection of continuously evolving data.

These tools monitor data movement to discover missing, partial, duplicated, and delayed data; identify changes in the content of spatiotemporal streams; and pinpoint anomaly hotspots based on persistence, pervasiveness, and priority. We conclude with lessons from FIT relevant to big data quality that are cause for optimism.

Short Bio:

Divesh Srivastava is the Head of Database Research at AT&T. He is a Fellow of the Association for Computing Machinery (ACM), the Vice President of the VLDB Endowment, on the Board of Directors of the Computing Research Association (CRA), on the ACM Publications Board and an associate editor of the ACM Transactions on Data Science (TDS). He has served as the managing editor of the Proceedings of the VLDB Endowment (PVLDB), as associate editor of the ACM Transactions on Database Systems (TODS), and as associate Editor-in-Chief of the IEEE Transactions on Knowledge and Data Engineering (TKDE). He has presented keynote talks at several international conferences, and his research interests and publications span a variety of topics in data management. He received his Ph.D. from the University of Wisconsin, Madison, USA, and his Bachelor of Technology from the Indian Institute of Technology, Bombay, India.

Conference Paper Presentations

Paper Sessions

S1	S1: Systems for High Performance ML and Numerical Computing (3 full + 3 short)	
	BigD355 "Towards Self-Tuning Parameter Servers"	Chris Liu, Pengfei Zhang, Bo Tang, Hang Shen, Ziliang Lai, Eric Lo, and Korris Chung (full)
	BigD214 "Massively Parallel Random Number Generation"	Christian Böhm and Claudia Plant (short)
	BigD598 "Extending the R Language with a Scalable Matrix Summarization Operator"	Sikder Tahsin Al Amin, Siva Uday Sampreeth Chebolu, and Carlos Ordonez (short)
	BigD228 "Improved Data Locality Using Morton-order Curve on the Example of LU Decomposition"	Martin Perdacher, Claudia Plant, and Christian Böhm (full)
	BigD512 "OverSketched Newton: Fast Convex Optimization for Serverless Systems"	Vipul Gupta, Swanand Kadhe, Thomas Courtade, Michael Mahoney, and Kannan Ramchandran (full)
	BigD438 "An Adaptive Layer Expansion Algorithm for Efficient Training of Deep Neural Networks"	Leo Chen, Pangfeng Liu, and Jan-Jan Wu (short)
S2	S2: Workflows and Monitoring (2 full + 5 short)	
	BigD633 "Effectively Using Remote I/O For Work Composition in Distributed Workflows"	Ryan Friese, Burcu Mutlu, Nathan Tallent, Joshua Suetterlein, and Jan Strube (short)
	BigD412 "Highly-Scalable Container Integrity Monitoring for Large-Scale Kubernetes Cluster"	Hirokuni Kitahara, Kugamoorthy Gajananan, and Yuji Watanabe (short)
	BigD679 "Anomaly Detection in Edge Nodes using Sparsity Profile"	Aekyeung Moon, Xiaoyan Zhuo, Jialing Zhang, Seung Woo Son, and Yun Jeong Song (full)
	BigD346 "Enhancing Open-Set Recognition using Clustering-based Extreme Value Machine (C-EVM)"	James Henrydoss, Steve Cruz, Chunchun Li, Manuel Gunther, and Terrance E Boulton (short)
	BigD562 "Chiron: Optimizing Fault Tolerance in QoS-aware Distributed Stream Processing Jobs"	Morgan Geldenhuys, Lauritz Thamsen, and Odej Kao (short)
	BigD646 "PB&J - Easy Automation of Data Science/Machine Learning Workflows"	Buvaneswari Ramanan, Lawrence Drabeck, Thomas Woo, Troy Cauble, and Anil Rana (full)
S3	S3: Big Human-Derived Data (3 full + 3 short)	
	BigD267 "T-EGAT: A Temporal Edge Enhanced Graph Attention Network for Tax Evasion Detection"	Yiyang Wang, Qinghua Zheng, Jianfei Ruan, Yuda Gao, Yan Chen, Xuanya Li, and Bo Dong (short)
	BigD747 "Knowledge Graph Enhanced Event Extraction in Financial Documents"	Kaihao Guo, Tianpei Jiang, and Haipeng Zhang (short)
	BigD379 "BELT: A Pipeline for Stock Price Prediction Using News"	Yingzhe Dong, Da Yan, Abdullateef Ibrahim Almudaifer, Sibo Yan, Zhe Jiang, and Yang Zhou (full)
	BigD245 "Real-time Karaoke Recommendations : Session-based Multi-Task Recommendations with Multivariate RNNs"	Shigeki Tanaka, Ryoki Wakamoto, and Yusuke Fukazawa (short)
	BigD742 "EPIC30M: An Epidemics Corpus Of Over 30 Million Relevant Tweets"	Junhua Liu, Trisha Singhal, Lucienne T.M. Blessing, Kristin L. Wood, and Kwan Hui Lim (full)
	BigD243 "Evaluating the Accuracy of Cloud NLP Services Using Ground-Truth Experiments"	Frank Pallas, Dimitri Staufner, and Jörn Kuhlentkamp (full)
S4	S4: Big Transportation Data (3 full + 3 short)	
	BigD376 "Learn to Earn: Enabling Coordination Within a Ride-Hailing Fleet"	Harshal Chaudhari, John Byers, and Evimaria Terzi (full)
	BigD362 "Knowledge- and Data-driven Services for Energy Systems using Graph Neural Networks"	Francesco Fusco, Bradley Eck, Robert Gormally, Mark Purcell, and Seshu Tirupathi (short)
	BigD331 "Crash prediction for a French highway network with an XAI-informed Bayesian hierarchical model"	Thomas Veran, Pierre-Edouard Portier, and François Fouquet (reg)
	BigD328 "Identifying Failing Point Machines from Sensor-Free Train System Logs"	Ying Yang, Xin Lou, Binbin Chen, Marianne Winslett, and Zbigniew Kalbarczyk (short)
	BigD327 "Machine-learning-based People-flow Simulation for Facility Layout Planning"	Satoshi Kuwamoto, Yu Kitano, and Akinori Asahara (short)
S5	BigD283 "Sequence Matching with Discriminative Binary Features for Robust and Fast Light-Rail Localization at High Frame Rate"	TINGXIAN WANG, KEBIN JIA, and MENG YAO (full)
	S5: Computational Models for Big Data (5 full)	

	BigD321 "Applications of Particle Swarm Optimization to System Identification and Supervised Learning"	Noah Schwalb and Edward Schwalb (full)
	BigD345 "State Action Separable Reinforcement Learning"	Ziyao Zhang, Liang Ma, Kin K. Leung, Konstantinos Poularakis, and Mudhakar Srivatsa (full)
	BigD369 "Asynchronous Online Federated Learning for Edge Devices with Non-IID Data"	Yujing Chen, Yue Ning, Martin Slawski, and Huzefa Rangwala (full)
	BigD456 "An Adversarial Domain Separation Framework for Septic Shock Early Prediction Across EHR Systems"	Farzaneh Khoshnevisan and Min Chi (full)
	BigD565 "Improving Model Training by Periodic Sampling over Weight Distributions"	Samarth Tripathi, Jiayi Liu, Saptik Dhar, Unmesh Kurup, and Mohak Shah (full)
S6	S6: Data Analytics Engines (3 full + 3 short)	
	BigD325 "HeAT - a Distributed and GPU-accelerated Tensor Framework for Data Analytics"	Markus Goetz, Charlotte Debus, Daniel Coquelin, Kai Krajsek, Claudia Comito, Philipp Knechtges, Bjoern Hagemeier, Michael Tarnawa, Simon Hanselmann, Martin Siggel, Achim Basermann, and Achim Streit (full)
	BigD258 "A Shared Memory Cache Layer across Multiple Executors in Apache Spark"	Wei Rang, Donglin Yang, and Dazhao Cheng (short)
	BigD408 "Extensible Data Skipping"	Paula Ta-Shma, Guy Khazma, Gal Lushi, and Oshrit Feder (full)
	BigD569 "Translation of Array-Based Loops to Spark SQL"	Md Hasanuzzaman Noor and Leonidas Fegaras (Short)
	BigD619 "A Genetic Optimization Physical Planner for Big Data Warehouses"	Soumia Benkrid, Yacine Mestoui, Ladjel Bellatreche, and Carlos Ordóñez (short)
	BigD529 "Accelerating the Configuration Tuning of Big Data Analytics with Similarity-aware Multitask Bayesian Optimization"	Ayat Fekry, Lucian Carata, Thomas Pasquier, and Andrew Rice (full)
S7	S7: Big Data Systems (3 full + 3 short)	
	BigD654 "BEAD: Batched Evaluation of Iterative Graph Queries with Evolving Analytics Demands"	Abbas Mazloumi, Chengshuo Xu, Zhijia Zhao, and Rajiv Gupta (short)
	BigD745 "HyGN: Hybrid Graph Engine for NUMA"	Tanuj Aasawat, Tahsin Reza, Kazuki Yoshizoe, and Matei Ripeanu (short)
	BigD211 "ReSpark: Automatic Caching for Iterative Applications in Apache Spark"	Michael Mior and Kenneth Salem (full)
	BigD415 "DSCEP: An Infrastructure for Decentralized Semantic Complex Event Processing"	Vitor Almeida, Sukanya Bhowmik, Guilherme Lima, Markus Endler, and Kurt Rothermel (short)
	BigD514 "Large-Scale Intelligent Microservices"	Mark Hamilton, Nick Gonsalves, Christina Lee, Anand Raman, Brendan Walsh, Siddhartha Prasad, Dalitso Banda, Lucy Zhang, Lei Zhang, and William Freeman (full, America)
	BigD548 "HReplca: A Dynamic Data Replication Engine with Adaptive Compression for Multi-Tiered Storage"	Hariharan Devarajan, Anthony Kougkas, and Xian-He Sun (full)
S8	S8: Big Biological and Physical Data (2 full + 5 short)	
	BigD620 "Combined Convolutional and Recurrent Neural Networks for Hierarchical Classification of Images"	Jaehoon Koo, Diego Klabjan, and Jean Utke (short)
	BigD607 "Graph Neural Network Architecture Search for Molecular Property Prediction"	Shengli Jiang and Prasanna Balaprakash (short)
	BigD464 "Gaussian Mixture Graphical Lasso with Application to Edge Detection in Brain Networks"	Hang Yin, Xinyue Liu, and Xiangnan Kong (short)
	BigD678 "Distributed de novo assembler for large-scale long-read datasets"	Sayan Goswami, Kisung Lee, and Seung-Jong Park (full)
	BigD698 "Deep Domain Adaptation based Cloud Type Detection using Active and Passive Satellite Data"	Xin Huang, Sahara Ali, Chenxi Wang, Zeyu Ning, Sanjay Purushotham, Jianwu Wang, and Zhibo Zhang (short)
	BigD713 "Structure Prediction from Neutron Scattering Profiles: A Data Sciences Approach"	Cristina Garcia-Cardona, Ramakrishnan Kannan, Travis Johnston, Thomas Proffen, and Sudip Seal (full)
	BigD719 "Deep Learning for Surface Wave Identification in Distributed Acoustic Sensing Data"	Vincent Dumont, Verónica Rodríguez Tribaldos, Jonathan Ajo-Franklin, and Kesheng Wu (short)
S9	S9: Big Human-Derived Data (3 full + 3 short)	

	BigD592 "Skill-based Career Path Modeling and Recommendation"	Aritra Ghosh, Beverly Woolf, Shlomo Zilberstein, and Andrew Lan (full)
	BigD491 "Analysis of Human Trafficking in North Carolina Based on Criminal Records: A Framework to Measure Human Trafficking Trends"	Shafie Gholizadeh, Matthew Phillips, Maryam Tavakoli Hosseinabadi, Daisy Leon, and James Rozier (short)
	BigD465 "Identifying FinTech Innovations Using BERT"	Doina Caragea, Mark Chen, Theodor Cojoianu, Mihai Dobri, Kyle Glandt, and George Mihaila (full)
	BigD411 "Valuing Player Actions in Counter-Strike: Global Offensive"	Peter Xenopoulos, Harish Doraiswamy, and Claudio Silva (full)
	BigD334 "ThanosNet: A Novel Trash Classification Method Using Metadata"	Alan Sun and Harry Xiao (short)
	BigD394 "MLAS: Metric Learning on Attributed Sequences"	Zhongfang Zhuang, Xiangnan Kong, Elke Rundensteiner, Jihane Zouaoui, and Aditya Arora (short)
S10	S10: Big Transportation and Spatial Data (3 full + 3 short)	
	BigD589 "A Simulation-based Approach for Large-scale Evacuation Planning"	Kazi Ashik Islam, Madhav Marathe, Henning Mortveit, Samarth Swarup, and Anil Vullikanti (short)
	BigD638 "Creating Realistic Power Distribution Networks using Interdependent Road Infrastructure"	Rounak Meyur, Madhav Marathe, Anil Vullikanti, Henning Mortveit, Samarth Swarup, Virgilio Centeno, and Arun Phadke (full)
	BigD418 "Leveraging an Efficient and Semantic Location Embedding to Seek New Ports of Bike Share Services"	Yuan Wang, Chenwei Wang, Yinan Ling, Keita Yokoyama, Hsin-Tai Wu, and Yi Fang (full)
	BigD409 "Dynamic Dispatching for Large-Scale Heterogeneous Fleet via Multi-agent Deep Reinforcement Learning"	Chi Zhang, Philip Odonkor, Shuai Zheng, Hamed Khorasgani, Susumu Serita, Chetan Gupta, and Haiyan Wang (short)
	BigD503 "Comprehensive Comparison of LSM Architectures for Spatial Data"	Qizhong Mao, Mohiuddin Abdul Qader, and Vagelis Hristidis (short)
	BigD494 "Lightweight, Embeddings Based Storage and Model Construction Over Satellite Data Collections"	Kevin Bruhwiler, Paahuni Khandelwal, Daniel Rammer, Samuel Armstrong, Sangmi Lee Pallickara, and Shrideep Pallickara (full)
S11	S11: Medical Data (3 full + 3 short)	
	BigD744 "Generating Realistic COVID19 X-rays with a Mean Teacher + Transfer Learning GAN"	Sumeet Menon, Joshua Galita, David Chapman, Aryya Gangopadhyay, Jayalakshmi Mangalagiri, Phuong Nguyen, Yaacov Yesha, Yelena Yesha, Babak Saboury, and Michael Morris (full)
	BigD567 "Heterogeneous Similarity Graph Neural Network on Electronic Health Records"	Zheng Liu, Xiaohan Li, Lifang He, Hao Peng, and Philip Yu (full)
	BigD586 "MuLan: Multilevel Language-based Representation Learning for Disease Progression Modeling"	Hyunwoo Sohn, Kyungjin Park, and Min Chi (full)
	BigD657 "From 5Vs to 6Cs: Operationalizing Epidemic Data Management with COVID-19 Surveillance"	Akhil Sai Peddireddy, Dawen Xie, Pramod Patil, Mandy Wilson, Dustin Machi, Srinivasan Venkatramanan, Brian Klahn, Przemyslaw Porebski, Parantapa Bhattacharya, Shirish Dumbre, Erin Raymond, and Madhav Marathe (short)
	BigD515 "r-LSTM: Time Series Forecasting for COVID-19 Confirmed Cases with LSTM-based Framework"	Mohammad Masum, Hossain Shahriar, Hisham Haddad, and Md. Shafiul Alam (short)
	BigD384 "Patient ADE Risk Prediction through Hierarchical Time-Aware Neural Network Using Claim Codes"	Jinhe Shi, Xiangyu Gao, Chenyu Ha, Yage Wang, Guodong Gao, and Yi Chen (short)
S12	S12: Theoretical Models for Big Data (3 full + 3 short)	
	BigD273 "Connecting MapReduce Computations to Realistic Machine Models"	Peter Sanders (full)
	BigD493 "A theoretical analysis of graph evolution caused by triadic closure and algorithmic implications"	Sara Ahmadian and Shahrzad Haddadan (full)
	BigD708 "Augmenting Deep Learning with Relational Knowledge from Markov Logic Networks"	Mohammad Maminur Islam, Somdeb Sarkhel, and Deepak Venugopal (full)
	BigD641 "BurstPU: Classification of Weakly Labeled Datasets with Sequential Bias"	Walter Gerych, Luke Buquicchio, Kevin Chandrasekaran, Abdulaziz Alajaji, Hamid Mansoor, Aidan Murphy, Elke Rundensteiner, and Emmanuel Agu (short)

	BigD645 "Understanding and Detecting Convergence for Stochastic Gradient Descent with Momentum"	Jerry Chee and Ping Li (short, Ameica)
	BigD695 "Neural Network Training Techniques Regularize Optimization Trajectory: An Empirical Study"	Cheng Chen, Junjie Yang, and Yi Zhou (short)
S13	S13: Health and Human Motion Data (3 full + 3 short)	
	BigD371 "Automated Machine Learning for the Classification of Normal and Abnormal Electromyography Data"	Marios Kefalas, Milan Koch, Victor Geraedts, Hao Wang, Martijn Tannemaat, and Thomas Bäck (full)
	BigD242 "Prognostication and Outcome-specific Risk Factor Identification for Diabetes Care via Private-shared Multi-task Learning"	Bin Liu, Ying Li, and Kenney Ng (full)
	BigD726 "NeuroKube: An Automated and Autoscaling Neuroimaging Reconstruction Framework using Cloud Native Computing and A.I."	Matthew Madany, Kyle Marcus, Steven Peltier, Mark H. Ellisman, and Ilkay Altintas (full)
	BigD281 "SMART: Emerging Activity Recognition with Limited Data for Multi-modal Wearable Sensing"	Madhuri Ghorpade, Haiquan Chen, Yuhong Liu, and Zhe Jiang (short)
	BigD315 "Segmentation and Recognition of Eating Gestures from Wrist Motion using Deep Learning"	Yadnyesh Luktuke and Adam Hoover (short)
	BigD422 "MASTGN: Multi-Attention Spatio-Temporal Graph Networks for Air Pollution Prediction"	peijiang zhao and Koji Zettsu (short)
S14	S14: Theoretical and Computational Models for Big Data (2 full + 5 short)	
	BigD570 "Combining Multiple Implicit-Explicit Interactions for Regression Analysis"	Minkyu Kim, Suan Lee, and Jinho Kim (full)
	BigD274 "'To Tell You the Truth' by Interval-Private Data"	Jie Ding and Bangjun Ding (full)
	BigD397 "Combining Global and Sequential Patterns for Multivariate Time Series Forecasting"	Zhaoxi Li, Jun He, Hongyan Liu, and Xiaoyong Du (short)
	BigD517 "Robust Adversarial Active Learning with a Novel Diversity Constraint"	Chengbin Sun, Hailong Sun, and Xudong Liu (short)
	BigD224 "Scale-Aware Graph-Based Machine Learning for Accurate Molecular Property Prediction"	Gyoung S. Na, Hyun Woo Kim, and Hyunju Chang (short)
	BigD459 "The Canonical Interval Forest (CIF) Classifier for Time Series Classification"	Matthew Middlehurst, James Large, and Anthony Bagnall (short)
	BigD672 "A Predictive Analytics Framework for Insider Trading Events"	Taruna Seth and Vipin Chaudhary (short)
S15	S15: Computational Models for Big Data (2 full + 5 short)	
	BigD651 "TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks"	Alexander Geiger, Dongyu Liu, Sarah Alnegheimish, Alfredo Cuesta-Infante, and Kalyan Veeramachaneni (full)
	BigD738 "GEDet: Adversarially Learned Few-shot Detection of Erroneous Nodes in Graphs"	Sheng Guan, Peng Lin, Hanchao Ma, and Yinghui Wu (full)
	BigD452 "Forecasting with Multiple Seasonality"	Tianyang Xie and Jie Ding (short)
	BigD693 "General Domain Adaptation Through Proportional Progressive Pseudo Labeling"	Mohammad Hashemi and Eric Keller (short)
	BigD749 "Fast and Accurate Tensor Decomposition without a High Performance Computing Machine"	Huan He, YuanZhe Xi, and Joyce C Ho (short)
	BigD354 "Consensus Variational and Monte Carlo Algorithms for Bayesian Nonparametric Clustering"	Yang Ni, David Jones, and Zeya Wang (short)
	BigD370 "Optimization of Graph Neural Networks with Natural Gradient Descent"	Mohammad Rasool Izadi, Yihao Fang, Robert Stevenson, and Lizhen Lin (short)
S16	S16: Efficient and Scalable Data Management and Mining (3 full + 3 short)	
	BigD402 "On Efficiently Processing Business Lineage Queries"	Himanshu Gupta, Rajmohan C, Sameep Mehta, and Kiran Pulapa (full)
	BigD617 "PYHESSIAN: Neural Networks Through the Lens of the Hessian"	Zhewei Yao, Amir Gholami, Kurt Keutzer, and Michael Mahoney (full)
	BigD702 "On the Mining of the Minimal Set of Time Series Data Shapelets"	Soukaina Filali Boubrahimi, Shah Muhammad Hamdi, Ruizhe Ma, and Rafal Angryk (full)
	BigD701 "Efficient Constrained Subgraph Extraction for Exploratory Discovery in Large Knowledge Graphs"	Sidan Gao, Nodirbek Korchiev, Vodelina Samatova, and Kemafor Anyanwu (short)
	BigD386 "Sketches: Fast Membership Scans for Continuous Variable Predicate Workloads."	Edward Schwalb (short)
	BigD445 "A new heuristic algorithm for fast k-segmentation"	Sabarish Vadarevu and Vijay Karamcheti (short)
S17	S17: Multi-structured Data - Big Variety Data (3 full + 3 short)	
	BigD622 "Raptor Zonal Statistics : Fully Distributed Zonal Statistics of Big Raster + Vector Data"	Samriddhi Singla and Ahmed Eldawy (full)

	BigD364 "A Regularized Model to Trade-off between Accuracy and Diversity in a News Recommender System"	Shaina Raza and Chen Ding (full)
	BigD510 "ExNav: An Interactive Big Data Exploration Framework for Big Unstructured Data"	Xiaoyu Ge, Xiaozhong Zhang, and Panos Chrysanthis (full)
	BigD332 "ExFaux: A Weakly Supervised Approach to Explainable Fauxtography Detection"	Ziyi Kou, Daniel Zhang, Lanyu Shang, and Dong Wang (short)
	BigD420 "Topic-Centric Unsupervised Multi-Document Summarization of Scientific and News Articles"	Amanuel Alambo, Cori Lohstroh, Erik Madaus, Swati Padhee, Brandy Foster, Tanvi Banerjee, Krishnaprasad Thirunarayan, and Michael Raymer (short)
	BigD468 "Semi-Supervised Learning and Feature Fusion for Multi-View Data Clustering"	Hadi Salman and Justin Zhan (short)
S18	S18: Data Streams and Time Series (3 full + 3 short)	
	BigD278 "C-SMOTE: Continuous Synthetic Minority Oversampling for Evolving Data Streams"	Alessio Bernardo, Heitor Murilo Gomes, Jacob Montiel, Bernhard Pfahringer, Albert Bifet, and Emanuele Della Valle (full)
	BigD294 "Efficient Record Linkage in Data Streams"	Dimitrios Karapiperis, Aris Gkoulalas-Divanis, and Vassilios Verykios (full)
	BigD449 "Unsupervised Drift Detection on High-speed Data Streams"	Vinicius Souza, Farhan Chowdhury, and Abdullah Mueen (full)
	BigD526 "StreamTeam-Football: Analyzing Football Matches in Real-Time on the Basis of Position Streams"	Lukas Probst, Heiko Schuldt, Philipp Seidenschwarz, and Martin Rumo (short)
	BigD629 "AutoML for Stream k-Nearest Neighbors Classification"	Maroua Bahri, Bruno Veloso, Albert Bifet, and Joao Gama (short)
	BigD471 "A Non-linear Function-on-Function Model for Regression with Time Series Data"	Qiyao Wang, Haiyan Wang, Chetan Gupta, Aniruddha Rajendra Rao, and Hamed Khorasgani (short)
S19	S19: Web Search and Mining (3 full + 3 short)	
	BigD419 "Analyzing Web Search Behavior for Software Engineering Tasks"	Nikitha Rao, Chetan Bansal, Thomas Zimmermann, Ahmed Hassan Awadallah, and Nachiappan Nagappan (full)
	BigD257 "Communication-Efficient Local Stochastic Gradient Descent for Scalable Deep Learning"	Sunwoo Lee, Qiao Kang, Ankit Agrawal, Alok Choudhary, and Wei-keng Liao (full)
	BigD454 "AML-SVM: Adaptive Multilevel Learning with Support Vector Machines"	Ehsan Sadrfaridpour, Korey Palmer, and Ilya Safto (full)
	BigD655 "Sketch and Scale: Geo-distributed tSNE and UMAP"	Viska Wei, Nikita Ivkin, Vladimir Braverman, and Alexander Szalay (short)
	BigD560 "Leveraging Natural Language Processing to Mine Issues on Twitter During the COVID-19 Pandemic"	Ankita Agarwal, Preetham Salehndam, Swati Padhee, William L. Romine, and Tanvi Banerjee (short)
	BigD482 "A Simple, Effective and Extendible Approach to Deep Multi-task Learning"	Yang Gao, Yi-Fan Li, Yu Lin, Hemeng Tao, and Latifur Khan (short)
S20	S20: Recommendation and Graph Mining (5 full)	
	BigD232 "Basket Recommendation with Multi-Intent Translation Graph Neural Network"	Zhiwei Liu, Xiaohan Li, Ziwei Fan, Stephen Guo, Kannan Achan, and Philip S. Yu (full)
	BigD230 "Graph Clustering with Embedding Propagation"	Carl Yang, Liyuan Liu, Mengxiong Liu, Zongyi Wang, Chao Zhang, and Jiawei Han (full)
	BigD255 "Ranking on Network of Heterogeneous Information Networks"	Zhe Xu, Si Zhang, Yinglong Xia, Liang Xiong, and Hanghang Tong (full)
	BigD373 "Unsupervised Multiple Network Alignment with Multinomial GAN and Variational Inference"	Yang Zhou, Jiaxiang Ren, Ruoming Jin, Zijie Zhang, Dejing Dou, and Da Yan (full)
	BigD740 "Mapping Network States using Connectivity Queries"	Alexander Rodriguez, Bijaya Adhikari, Andres D. Gonzalez, Charles Nicholson, Anil Vullikanti, and B. Aditya Prakash (full)
S21	S21: Data Preprocessing and Integration (5 full)	
	BigD731 "GLIMA: Global and Local Time Series Imputation with Multi-directional Attention Learning"	Qiuling Suo, Weida Zhong, Guangxu Xun, Jianhui Sun, Changyou Chen, and Aidong Zhang (full)
	BigD485 "Supervised Topic Compositional Neural Language Model for Clinical Narrative Understanding"	Xiao Qin, Cao Xiao, Tengfei Ma, Tabassum Kakar, Susmitha Wunnava, Xiangnan Kong, Elke Rundensteiner, and Fei Wang (full)
	BigD260 "Pattern-enhanced Named Entity Recognition with Distant Supervision"	Xuan Wang, Yingjun Guan, Yu Zhang, Qi Li, and Jiawei Han (full)

	BigD261 "Textual Evidence Mining via Spherical Heterogeneous Information Network Embedding"	Xuan Wang, Yu Zhang, Aabhas Chauhan, Qi Li, and Jiawei Han (full)
	BigD301 "Dynamic Chunkwise CNN for Distantly Supervised Relation Extraction"	Fangbing Liu and Qing Wang (full, Australia)
S22	S22: Big Variety Data (3 full + 3 short)	
	BigD341 "App-Aware Response Synthesis for User Reviews"	Umar Farooq, A.B. Siddique, Fuad Jamour, Zhijia Zhao, and Vagelis Hristidis (full)
	BigD668 "The Role of Citizens in Countering Misinformation: A Case Study of COVID-19 Infodemic"	Nicholas Micallef, Bing He, Srijan Kumar, Mustaque Ahamad, and Nasir Memon (full)
	BigD324 "VizObj2Vec: Contextual Representation Learning for Visual Objects in Video-frames"	Ahnaf Farhan and M. Shahriar Hossain (full)
	BigD649 "Instance Matching in Knowledge Graphs Through Dynamic, Distributed and Affinity-Preserving Random Walk"	Ali Assi, Mohamed Elati, and Wajdi Dhifli (short)
	BigD488 "Piaget: A Probabilistic Inference Approach for Geolocating Historical Buildings"	Sasan Tavakkol, Cyrus Shahabi, Feng Han, and Raimondas Kiveris (short)
	BigD568 "A Proficient Spelling Analysis Method Applied to Herbal and Dietary Supplement Discovery in a Large Clinical Corpus"	Terri Workman, Yijun Shao, Guy Divita, and Qing Zeng-Treitler (short)
S23	S23: Threat Detection and Web Search (2 full + 5 short)	
	BigD316 "FairFL: A Fair Federated Learning Approach to Reducing Demographic Bias in Privacy-Sensitive Classification Models"	Daniel Zhang, Ziyi Kou, and Dong Wang (full)
	BigD552 "Towards an Open Format for Scalable System Telemetry"	Teryl Taylor, Frederico Araujo, and Xiaokui Shu (full)
	BigD351 "Detecting and Mitigating Adversarial Attacks Through Robust Feature Alignment"	Scott Freitas, Shang-Tse Chen, Zijie Wang, and Duen Horng Chau (short)
	BigD724 "Gaussian Function On Response Surface Estimation"	Mohammadhossein Toutiaee and John Miller (short)
	BigD393 "MBead: Semi-supervised Multilabel Behaviour Anomaly Detection on Multivariate Temporal Sensory Data"	Suwen Lin, Louis Faust, Sidney D'Mello, Gonzalo Martinez, and Nitesh Chawla (short)
	BigD307 "TabSim: A Siamese Neural Network for Accurate Estimation of Table Similarity"	Maryam Habibi, Johannes Starlinger, and Ulf Leser (short)
	BigD431 "A Hybrid Deep Model for Learning to Rank Data Tables"	Mohamed Trabelsi, Zhiyu Chen, Brian D. Davison, and Jeff Hefflin (short)
S24	S24: Privacy Preserving Big Data Collection/Analytics (4 full + 1 short)	
	BigD489 "A Study of Self-Privacy Violations in Online Public Discourse"	Prasanna Umar, Anna Squicciarini, and Sarah Rajtmajer (full)
	BigD497 "Privacy-Aware Federated Learning for Page Recommendation"	Shuai Zhao, Roshani Bharati, Cristian Borcea, and Yi Chen (full)
	BigD544 "Stochastic Adaptive Line Search for Differentially Private Optimization"	Chen Chen and Jaewoo Lee (full)
	BigD687 "Broadening Differential Privacy for Deep Learning Against Model Inversion Attacks"	Qiuchen Zhang, Jing Ma, Yonghui Xiao, Jian Lou, and Li Xiong (full)
	BigD689 "Towards Training Robust Private Aggregation of Teacher Ensembles Under Noisy Labels"	Qiuchen Zhang, Jing Ma, Jian Lou, Li Xiong, and Xiaoqian Jiang (short)
S25	S25: Recommender Systems and Graphs (3 full + 3 short)	
	BigD353 "An Evaluation of Backpropagation Interpretability for Graph Classification with Deep Learning"	Kenneth Teo Tian Shun, Eko Edita Limanta, and Arijit Khan (full)
	BigD542 "FP-ExtVP: Accelerating Distributed SPARQL Queries by Exploiting Load-adaptive Partitioning"	Jiamin Lu, Cheng Yang, Bingfa Wang, and Jun Feng (full)
	BigD540 "Cost-Aware Influence Maximization in Multi-Attribute Networks"	Ioulia Litou and Vana Kalogeraki (full)
	BigD293 "Inventory Based Recommendation Algorithms"	Du Chen, Yuming Deng, Guangrui Ma, Hao Ge, Yunwei Qi, Ying Rong, Xun Zhang, and Huan Zheng (short)
	BigD451 "Self-supervised Hierarchical Graph Neural Network for Graph Representation"	Sambaran Bandyopadhyay, Manasvi Aggarwal, and M. Narasimha Murty (short)
	BigD627 "Massively Parallel Graph Drawing and Representation Learning"	Christian Böhm and Claudia Plant (short)
S26	S26: Graph Mining and Embedding (3 full + 3 short)	
	BigD400 "Hypergraph Attention Isomorphism Network by Learning Line Graph Expansion"	Sambaran Bandyopadhyay, Kishalay Das, and M. Narasimha Murty (full)
	BigD594 "TI-GCN: A Dynamic Network Embedding Method with Time Interval Information"	Yali Xiang, Yun Xiong, and Yangyong Zhu (full)

	BigD521 "A Comprehensive and Adversarial Approach to Unsupervised Embedding Learning"	Yi-Zhan Hsu, Sungwon Han, Sungwon Park, Meeyoung Cha, and Cheng-Te Li (full)
	BigD252 "Adaptive Community Search in Dynamic Networks"	Ioanna Tsalouchidou, Francesco Bonchi, and Baeza-Yates Ricardo (short)
	BigD358 "Efficient Computing of PageRank Scores on Exact Expected Transition Matrix of Large Uncertain Graph"	Takayasu Fushimi, Kazumi Saito, Kouzou Ohara, Masahiro Kimura, and Hiroshi Motoda (short)
	BigD360 "AutoAudit: Mining Accounting and Time-Evolving Graphs"	Meng-Chieh Lee, Yue Zhao, Aluna Wang, Pierre Jinghong Liang, Leman Akoglu, Vincent S. Tseng, and Christos Faloutsos (short)
S27	S27: Temporal and StreamData Mining (2 full + 5 short)	
	BigD450 "An Unsupervised Misinformation Detection Framework to Analyze the Users using COVID-19 Twitter Data"	Aarzoo Dhiman and Durga Toshniwal (full)
	BigD423 "Practical Range Counting over Data Streams"	Ran Bai, Ziliang Lai, Eric Lo, Wing-Kai Hon, and Pengfei Zhang (full)
	BigD361 "Co-Embedding Attributed Networks with External Knowledge"	Pei-Chi Lo and Ee-Peng Lim (short)
	BigD399 "Fusion-3DCNN-PFP: A dynamic system for discovering patterns of predicted congestion"	Minh-Son Dao, Ngoc-Thanh Nguyen, Rage Uday Kiran, and Koji Zettsu (short)
	BigD500 "Temporal Graph Neural Networks for Social Recommendation"	Ting Bai, Youjie Zhang, Bin Wu, and Jian-Yun Nie (short)
	BigD463 "Drift-Aware Multi-Memory Model for Imbalanced Data Streams"	Amir Abolfazli and Eirini Ntoutsi (short)
	BigD437 "Acceleration of large deep learning training with hybrid GPU memory management of swapping and re-computing"	Haruki Imai, Tung D. Le, Yasushi Negishi, and Kiyokuni Kawachiya (short)
S28	S28: Representation Learning (2 full + 5 short)	
	BigD368 "Learning Similarity-Preserving Meta-Embedding for Text Mining"	Jidapa Thadajarassiri, Cansu Sen, Thomas Hartvigsen, Xiangnan Kong, and Elke Rundensteiner (full)
	BigD572 "Automating Domain Squatting Detection Using Representation Learning"	Pablo Loyola, Kugamoorthy Gajananan, Hirokuni Kitahara, Yuji Watanabe, and Fumiko Satoh (full)
	BigD457 "MLANE: Meta-Learning Based Adaptive Network Embedding"	Chen Cui, Ning Yang, and Philip S. Yu (short)
	BigD477 "Characterizing and Utilizing the Interplay Between Core and Truss Decompositions"	Penghang Liu and Ahmet Erdem Sariyüce (short)
	BigD550 "A Hybrid Salient Object Detection with Global Context Awareness"	Minglin Hong, Xiaolin Li, Jing Wang, Haiyang He, and Shiguo Huang (short)
	BigD426 "GAMIT: A New Encoder-Decoder Framework with Graphical Space and Multi-grained Time for Traffic Predictions"	Zhixiang He, Chi-Yin Chow, and Jia-Dong Zhang (short)
	BigD635 "MOIRE: Mixed-Order Poisson Regression towards Fine-grained Urban Anomaly Detection at Nationwide Scale"	Masamichi Shimosaka, Kota Tsubouchi, Yanru Chen, Yoshiaki Ishihara, and Junichi Sato (short)

Industry and Government Paper Presentations

I&G 1: Big Data: Machine Learning Applications I	
N213	Search Wandering Score: Predicting Timings of Online Shopping based on Wandering in User's Web Search Queries
N229	Estimating Risk-Adjusted Hospital Performance
N240	Methodology for Interpretable Reinforcement Learning Model for HVAC Energy Control
N206	V-WorkGen: Virtual Workload Generation Tool for Connected Automotive Services
N207	Complementary Recommendations Using Deep Multi-modal Embeddings For Online Retail
I&G 2: Big Data: Machine Learning Applications II	
N239	Intervention Recommendation for Improving Disability Employment

N211	Dynamic Creative Optimization in Verizon Media Native Advertising
N233	Merchant Identity Recognition Using Credit Card Transactions
N248	Deep Neural Query Understanding System at Expedia Group
N225	2D-ATT: Causal Inference for Mobile Game Organic Installs with 2-Dimensional Attentional Neural Network
N239	Intervention Recommendation for Improving Disability Employment
I&G 3: Big Data: Feature and Anomaly Detection I	
N250	SOSNet: A Graph Convolutional Network Approach to Fine-Grained Cyberbullying Detection
N209	Deep Anomaly Detection with Ensemble-Based Active Learning
N220	Image segmentation for dust detection using semi-supervised machine learning
N247	An End-to-End Context Aware Anomaly Detection System
N253	FLOps: On Learning Important Time Series Features for Real-Valued Prediction
I&G 4: Big Data: Feature and Anomaly Detection II	
N203	Effective Detection of Rare Anomalies from Massive Waveform Data Using Heterogeneous Clustering
N242	CNN Application in Detection of Privileged Documents in Legal Document Review
N243	Building an Automated and Self-Aware Anomaly Detection System
N219	Doctor for Machines: A Failure Pattern Analysis Solution for Industry 4.0
N249	Temporal Tensor Transformation Network for Multivariate Time Series Prediction
I&G 5: Big Data ML Platforms and Systems I	
N204	EGAD: Evolving Graph Representation Learning with Self-Attention and Knowledge Distillation for Live Video Streaming Events
N236	Conception of a Reference Architecture for Machine Learning in the Process Industry
N256	The Danish National Energy Data Lake: Requirements, Technical Architecture, and Tool Selection
N218	Smart-ML: A System for Machine Learning Model Exploration using Pipeline Graph
N228	FDC Cache: Semantics-driven Federated Caching and Querying for Big Data
I&G 6: Big Data ML Platforms and Systems II	
N208	Ultron-AutoML: an open-source, distributed, scalable framework for efficient hyper-parameter optimization
N221	CANON: Complex Analytics of Network of Networks for Modeling Adversarial Activities
N215	PAIRS AutoGeo: An Automated Machine Learning Framework for Massive Geospatial Data
N212	Statistical Reasoning of Zero-Inflated Right-Skewed User-Generated Big Data A/B Testing
N251	A Verifiable Imputation Analysis for Univariate Time Series and Enabling Package
I&G 7: Data Quality and Privacy	
N210	Location YardStick: Calculation of the Location Data Value Depending on the Users' Context

N214 N237	Empirical Evaluation of Federated Learning with Local Privacy for Real-World Application The Missing Input Problem
N254	DQLearn: A Toolkit for Structured Data Quality Learning
N258	Privacy-Friendly ID-Free Digital Ad Targeting using URL Embeddings

Tutorials

TUTORIAL 1: Big Data Applications for High-Impact / Emerging Topics: Case Studies on Molecular Biology and Its Applications to COVID-19

Presenters: Ka-Chun Wong, Jiecong Lin, Shixiong Zhang, Xiangtao Li

Abstract:

The development of big data foundation, infrastructure, management, and algorithms have been demonstrated fruitful in the past years. To ensure broad impacts, the next step is to deploy those developed assets into real-life applications. In particular, the recent advances in high-throughput have revolutionized the field of molecular biology where there are high volumes of sequencing data with molecular veracity, functional variety, and real-life values. At the beginning (30 mins), we will teach you the basic concepts in molecular biology. Specifically, we will focus on the central dogma of molecular biology and the corresponding classic bioinformatics tools for extracting useful information from the high-throughput sequencing data at different stages in genetics. As the main content (1 hrs), we will show you different big data case studies for high-impact topics in molecular biology as published on high-impact journals (i.e. IF>5). Specifically, the fundamental topics in big data sequence modelling and pattern recognition for molecular biology (e.g. DNA motifs) will be introduced as the first case study. After that, we will switch to two advanced topics in molecular biology for personalized medicine solutions. The first one lies in the off-target predictions for CRISPR-Cas9 gene editing, while the second one is the early cancer detection from blood. Lastly (30 mins), we will demonstrate how different big data tools can be applied to COVID-19 in the context of molecular modelling (e.g. vaccine design) and text mining on social media.

TUTORIAL 2: Big Data System Benchmarking: State of the Art, Current Practices, and Open Challenges

Presenters: Chao Zhang, Jiaheng Lu

Abstract:

Big data system benchmarking enables practitioners and developers to assess the systems' functionality and performance so that they can make wise decision to choose the proper big data systems, or improve them. As we are witnessing the emergence and evolvement of various benchmarks for big data systems, either in the form of macrobenchmark or micro-benchmark, it is crucial to thoroughly study, analyze, and understand the key techniques and applications of those benchmarks. In this tutorial, we offer a comprehensive presentation of a wide range of the state-of-the-art benchmarks with a focus on big data systems. We classify these benchmarks into five categories: Map-Reduce based system benchmarking, SQL-based analytical system benchmarking, NoSQL-based database benchmarking, Big graph system benchmarking, and Multi-model database benchmarking. We discuss the key techniques of each approach, as well as the current practices. We also provide insights on the research challenges and directions for benchmarking different big data systems.

TUTORIAL 3: Systems and Algorithms for Massively Parallel Graph Mining

Presenters: Da Yan, Guimu Guo

Abstract:

Big graph processing systems such as Pregel, GraphLab, GraphX and Gemini have become increasingly popular thanks to their emphasis on ease of programming. Unfortunately, these frameworks are dominantly designed for IO-bound execution and are only suitable for problems with a low time complexity. Graph mining problems such as finding dense and frequent subgraph structures usually have a very high time complexity, and when IO-bound systems are applied, the performance is a catastrophe. However, this problem is still not getting enough attention among many graph-parallel algorithm and system researchers who are still using those IO-bound systems to address compute-heavy graph mining problems. In this tutorial, we explicitly categorize the popular graph mining problems into IO-heavy and CPU-heavy categories, and provide prior evidence that CPU-heavy graph mining problems should not be addressed using IO-bound systems which can lead to performance worse than even a serial algorithm. We then introduce two recent compute-intensive solutions to mining dense subgraph structures and frequent subgraph patterns, respectively, that satisfactorily address the IO-bound issue of existing systems. The key design is to expose an explicit task-based divide-and-conquer API to users, in contrast to the existing iterative computation paradigms. We will also show how to develop popular graph mining algorithms in these frameworks, including finding maximum cliques, triangle counting, finding maximal γ -quasi-cliques, and finding k-plexes.

TUTORIAL 4: Semantic Exploration of Big Data

Presenters: Maria Krommyda, Verena Kantere

Abstract:

As the volume of the available information increases, more and more datasets meet the criteria of being characterized as Big Data, thus creating the need for systems that can support their exploration. As the characteristics of Big Data create new challenges that render existing solutions obsolete, in the recent years there have been many efforts to identify new techniques and develop systems that will enable the semantic exploration of Big Data by a wider audience. Some of these systems rely on the Semantic Web standard, which allows

them to provide incremental navigation based on the semantic ontologies, while others take advantage of the structure of the RDF model to provide exploration based on graph topology. Other systems focus on the datasets offered through SPARQL endpoints either aiming to identify their schema or to provide support for query creation. The technique over which a system is designed is very important, as it determines the challenges that are addressed, the datasets and use cases that can be used as well as any limitations. In addition, specific design decision may create differentiation between systems implementing the same technique. We propose here, the presentation of a tutorial that will present the characteristics of the Big Data, discuss the challenges for their exploration and understanding, as well as techniques that aim to overcome them. The tutorial will offer to the audience a deep understanding of the strengths and weaknesses of the techniques, the use cases and datasets that can be applied to as well as an overview of the available systems currently implementing them and their functionalities.

TUTORIAL 5: A Tutorial on Topological Data Analysis in Text Mining

Presenters: Wlodek Zadrozny, Shafie Gholizadeh

Abstract:

Topological Data Analysis (TDA) introduces methods that capture the underlying structure of shapes in data. Despite the old history of computational geometry and computational topology in applied mathematics, utilization of topology in data science is relatively a new phenomenon. Within the last decade, TDA has been mostly examined in unsupervised machine learning tasks. As an alternative to the conventional algorithms, TDA has been often considered due to its capability to deal with high-dimensional data in different tasks including but not limited to clustering, dimensionality reduction or descriptive modeling. This tutorial will focus on applications of topological data analysis to text, and in particular to text classification. After introducing the fundamentals we will show three ways in which topological information can be added to improve the accuracy of classification. More specifically, we explain three different methods of extracting topological features from textual documents, using as the underlying representations of text the two most popular methods, namely term frequency vectors and word embeddings, and also without using any conventional features. In addition, we show how even the simplest out of the box topological methods can be used to provide similarity judgments, e.g. topological plots of classical novels.

TUTORIAL 6: Data Sources, Tools, and Techniques for Big Data-driven Machine Learning in Heliophysics

Presenters: Azim Ahmadzadeh, Berkay Aydin, Dustin J. Kempton, Rafal A. Angryk

Abstract:

During the past decade, Georgia State University's (GSU) Data Mining Lab (DMLab) has been conducting research on a wide range of topics centering on understanding, detection, and forecast of solar events, those of which can (directly or indirectly) have significant economic and collateral impacts on mankind, through electromagnetic radiation and energetic particles. The close collaboration of the Computer Scientists and Solar Physicists with the sole dedication to research on solar events using advanced statistical tools, machine learning (ML) and deep learning (DL), resulted in a couple of hundreds of in-depth studies in this domain. Many of these studies have been published in prestigious journals such as Nature's Scientific Data and The Astrophysical Journal. We would like to prepare a tutorial on some of the methodologies we engineered, the challenges we faced, and the products we put together. We believe our solutions and products can stimulate new data-driven discoveries in heliophysics, as well as to serve and inspire communities of other domains.

TUTORIAL 7: Big data stream mining

Presenters: Bartosz Krawczyk

Abstract:

The data revolution over the last two decades has changed almost every aspect of contemporary data analytics. One must consider the fact that the size of data is constantly growing, and one cannot store all of it. Data is in fast motion, constantly expanding and changing its properties. Velocity of data gave rise to the notion of data streams, potentially unbounded collections of data that continuously flood the system. As new data is continuously arriving, storing all the data stream is not a viable option. One needs to analyze new instances on-the-fly, incorporate the useful information into the classifier, and discard them. Data streams are also subject to a phenomenon known as concept drift, where the properties of stream are subject to unexpected changes over time. This includes not only the discriminatory power of features, but also the size of the feature space, ratios of instances among classes, as well the emergence and disappearance of features and classes. To accommodate such characteristics, data streams inspired the development of new families of algorithms capable of continuously integrating new data, while being robust to its evolving nature. This tutorial aims at introducing the audience to both basic concepts of learning from data streams, as well as to advanced contemporary research topics in this area. Special properties and requirements of data stream mining will be highlighted, to clearly emphasize the state of this field, as well as underlying challenges and emerging trends.

TUTORIAL 8: Learning from complex medical data

Presenters: Myra Spilopoulou, Panagiotis Papapetrou, Jaakko Hollmen

Abstract:

Big data analytics and machine learning methods are intensively employed in medicine and healthcare. Electronic Health Records (EHRs) are perceived as big patient data. On them, scientists strive to perform predictions on patients' progress, to understand and predict response to treatment, to detect adverse drug effects and factors for cardiovascular disease, and many other learning tasks. Medical researchers are also interested in learning from cohorts of population-based studies and of experiments. Learning tasks include the identification of disease predictors that can lead to new diagnostic tests and the acquisition of insights on interventions. In this tutorial, we elaborate on data sources, methods, and case studies in medical mining. Next to the aforementioned conventional data sources, we address the potential of data from mobile devices. We discuss the learning problems that can be solved with those data, we present case studies and investigate the methods needed to prepare and mine those data and to present the results to a medical expert. Furthermore, we will emphasize the need for interpretable and explainable models that can inspire trust and facilitate informed decision making. Towards this goal we will discuss and elaborate on actionable models for complex EHR data and their applicability on the interpretation of black-box models, such as deep learning architectures.

TUTORIAL 9: Big Sequence Management

Presenters: Karima Echihabi, Kostas Zoumpatians, Themis Palpanas

Abstract:

Data series are a prevalent data type that has attracted lots of interest in recent years. Specifically, there has been an explosive interest towards the analysis of large volumes of data series in many different domains. This is both in businesses (e.g., in mobile applications) and in sciences (e.g., in biology). In this tutorial, we focus on applications that produce massive collections of data series, and we provide the necessary background on data series storage, retrieval and analytics. We look at systems historically used to handle and mine data in the form of data series, as well as at the state of the art data series management systems that were recently proposed. Moreover, we discuss the need for fast similarity search for supporting data mining applications, and describe efficient similarity search techniques, indexes and query processing algorithms. Finally, we look at the gap of modern data series management systems in regards to support for efficient complex analytics, and we argue in favor of the integration of summarizations and indexes in modern data series management systems. We conclude with the challenges and open research problems in this domain.

Workshops

International Workshop on Big Data Analytics for Cyber Threat Hunting (CyberHunt 2020) Workshop Chair: Vasileios Mavroeidis PC Chair: Nils Gruschka PC Chair Assistant: Mateusz Zych		
Time (Eastern Time)	Title	Presenter/Author
9:00-9:05	Opening Remarks	Organizers
9:05-9:45	Keynote Presentation: The (new) attack surfaces of data-learned models - Adversarial attacks and defenses for ML models	Fabio Zennaro
9:45-10:10	CTI ANT: Hunting for Chinese Threat Intelligence (S01211)	Chia-En Tsai
10:10-10:35	Towards Fighting Cybercrime: Malicious URL Attack Type Detection using Multiclass Classification (S01212)	Tariro Manyumwa
10:35-11:00	Does Sophisticating Double Arbiter PUF Design Ensure its Security? Performance and Security Assessments on 5-1 DAPUF (S01207)	Meznah Alamro
11:00-11:25	An Approach to Combining Medical Device Fault Analysis with Trusted Computing Forensics (S01216)	Ian Oliver
Coffee Break 11:25-11:40		
11:40-12:05	The Applicability of Authorship Verification to Swedish Discussion Forums (S01209)	Lukas Lundmark
12:05-12:30	Firearm Detection via Convolutional Neural Networks: Comparing a Semantic Segmentation Model Against End-to-End Solutions (S01218)	Alexander Egiazarov
12:30-12:55	Assessment of Data Augmentation Techniques for Firearm Detection in Surveillance Videos (S01204)	Jacob Rose
12:55-13:20	Cyber Threat Hunting Through Automated Hypothesis and Multi-Criteria Decision Making (S01205)	Antonio Jose Horta Neto
13:20-13:25	Closing Remarks	Organizers

9th Workshop on Scalable Cloud Data Management <i>Felix Gessert (Baqend), Wolfram Wingerath (Baqend), Benjamin Wollmer (University of Hamburg), Norbert Ritter (University of Hamburg)</i>		
Time	Title	Presenter/Author
9:00 - 9:10 am	Opening Remarks	Felix Gessert, Norbert Ritter
9:10 - 9:55 am	Keynote TBA	TBA
9:55 - 10:25 am	Quasi-optimal Data Placement for Secure Multi-tenant Data Federation on the Cloud (BigD344)	Ji LIU (Big Data Lab, Baidu Inc., Beijing, China)
10:25 - 10:55 am	Building and Evaluation of Cloud Storage and Datasets Services on AI and HPC Converged Infrastructure (S03204)	Yusuke Tanimura (National Institute of Advanced Industrial Science and Technology Aomi 2-4-7, Koto-ku, Tokyo, Japan)
10:55 - 11:25 am	EvoBench - A Framework for Benchmarking Schema Evolution in NoSQL (S03203)	Mark Lukas Möller (Computer Science Institute University of Rostock)
11:25 - 11:45 am	Coffee Break	
11:45 - 12:15 pm	Operator as a Service: Stateful Serverless Complex Event Processing (S03205)	Manisha Luthra (Technical University of Darmstadt, Germany)
12:15 - 12:45 pm	Approximate Distance-based Anomaly Detection at Massive Scale (S03206)	Sivam Pasupathipillai (DISI, University of Trento - Trento, Italy)
12:45 - 1:15 pm	Bridging BAD Islands: Declarative Data Sharing at Scale (S03201)	Xikui Wang (Donald Bren School of Information and Computer Sciences University of California Irvine)
1:15 - 1:30 pm	Closing Remarks	

6th International Workshop to Improve Big Data Science Project Team Processes <i>Workshop Chairs: Jeffrey Saltz & Mary Magee Quinn</i> December 11th		
Time	Title	Presenter/Author
9:00 am	On the Large-scale Graph Data Processing for User Interface Testing in Big Data Science Projects	Yasin Uygun, Ramazan Faruk Oğuz, Erdi Olmezogullari, and Mehmet Siddik Aktaş
9:30am	Making "Magic" with Engineered Decisions, Data, and Processes: A Hospital Operations Center	John Celona, Louis Halamek, and Adam Seiver
10:00 am	Requirements Engineering Practices and Challenges in the Context of Big Data Software Development Projects: Early Insights from a Case Study	Darlan Arruda and Rodrigo Laigner
	Coffee Break	
10:40am	Maintaining NoSQL Database Quality During Conceptual Model Evolution	Pablo Suárez-Otero, Michael J. Mior, Maria José Suárez-Cabal, and Javier Tuya
11:10am	Adopting Agile Software Development Methodologies in Big Data Projects – a Systematic Literature Review of Experience Reports	Iva Krasteva and Sylvia Ilieva
11:40am	Recommended Methods for Using the 2020 NIST Principles for AI Explainability	Mary Quinn, Blake Piper, James Bliss, and David Keever
12:10pm	Exploring Frameworks Big Data Science Teams Use to Structure and Coordinate their Projects	Jeffrey Saltz and Nick Hotz
	Closing Remarks	

The 3rd International Workshop on “Big Data Engineering and Analytics in Cyber-Physical Systems (BigEACPS)”

Workshop Chairs: Akbar Namin

Time	Title	Presenter/Author
1:30 – 2:00	A Sensitivity Analysis of Evolutionary Algorithms in Generating Secure Configurations (Paper ID: S05202)	Shuvalaxmi Dass
2:00 – 2:30	Email Embeddings for Phishing Detection (Paper ID: S05205)	Luis Felipe Gutiérrez
2:30 – 3:00	Predicting Consequences of Cyber-Attacks (Paper ID: S05206)	Prerit Datta
3:00 – 3:30	Predicting Emotions Perceived from Sounds (Paper ID: S05203)	Faranak Abri
3:30 – 4:00	Ethereum Smart Contracts: Vulnerabilities and their Classification (Paper ID: S05201)	Zulfiqar Khan
4:00 – 4:30	A Concern Analysis of Federal Reserve Statements: The Great Recession vs. The COVID-19 Pandemic (Paper ID: S05204)	Luis Felipe Gutiérrez

5th Workshop on Real-time Stream Analytics, Stream Mining, CER/CEP & Stream Data Management in Big Data

Workshop Chairs: Sabri SKHIRI, Albert BIFET, Alessandro MARGARA

Time	Title	Presenter/Author
9:00 - 9:45	Keynote 1: The spectrum of stream processing with Apache Flink (S07204)	Till Rohrmann - Ververica
9:45 - 10:30	Keynote 2: Scaling Apache Pulsar Functions (S07205)	Sanjeev Kulkarni - Splunk
10:30 - 11:15	Keynote 3: How to architect data pipelines with Structured Streaming in Apache Spark (S07206)	Tathagata Das - Databricks
	Coffee Break	
11:25 - 11:40	Optimizing Convergence for Iterative Learning of ARIMA for Stationary Time Series (S07202)	Kevin Styp-Rekowski, Florian Schmidt, and Odej Kao
11:40 - 11:55	Extending Kafka Streams for Complex Event Recognition (S07203)	Samuele Langhi, Riccardo Tommasini, and Emanuele Della Valle
11:55 - 12:20	Streaming Time Series Forecasting using Multi-Target Regression with Dynamic Ensemble Selection (BigD605)	Dihia Boulegane, Albert Bifet, Haytham Elghazel, and Giyyarpuram Madhusudan
12:20 - 12:45	Flexible Executor Allocation without Latency Increase for Stream Processing in Apache Spark (BigD527)	Yuta Morisawa, Masaki Suzuki, and Takeshi Kitahara
	Coffee Break	
12:55 - 13:20	Smart Resource Management for Data Streaming using an Online Bin-packing Strategy (BigD469)	Oliver Stein, Ben Blamey, Johan Karlsson, Alan Sabirsh, Ola Spjuth, Andreas Hellander, and Salman Toor
13:20 - 14:45	Monitoring Networks with Queries Evaluated by Edge Computing (BigD671)	Quangtri Thai, Ordóñez Carlos, and Omprakash Gnawali,
14:45 - 15:10	HerdMonitor: Monitoring Live Migrating Container Resource and Performance Metrics in Cloud Environments (BigD746)	Alejandro Gonzalez and Emmanuel Arzuaga
	Closing Remarks	

The Fourth Annual Workshop on Applications of Artificial Intelligence in the Legal Industry <i>Workshop Chairs: Jianping Zhang, Nathaniel Huber-Fliflet, and Haozhen Zhao</i>			
Time (EST)	Title	Speaker/Authors	Country
9:30-9:35	Opening Remarks		
9:35-10:00	Invited Talk: Data and “Document” Trends in the Legal Industry	Robert Keeling - Sidley Austin LLP	
10:00-10:20	S06210 “SALKG: A Semantic Annotation System for Building a High-quality Legal Knowledge Graph”	Mingwei Tang, Cui Su, Haihua Chen, Jingye Qu, and Junhua Ding	China
10:20-10:40	S06205 “AI to Support Public Administration: the Case of the Analysis of Italian Local Taxation Documents”	Antonio Ballarin, Marco Brambilla, Emre Calisir, Giancarlo Caputo, Edgardo Carena, Andrea De Lullo, Andrea Mauri, Fabrizio Minuti, Israel Salgado, and Gianluca Sironi	Italy
10:40-11:00	S06206 “Preventing litigation with a predictive model of COVID-19 ICUs occupancy”	Chiara Gallese, Elena Falletti, Marco S. Nobile, Lucrezia Ferrario, Fabrizio Schettini, and Emanuela Foglia	Italy
11:00-11:20	S06212 “Extracting Key Elements from Legal Permits: A Use Case in the Environmental Domain”	Anna Breit, Laura Waltersdorfer, Fajar Ekaputra, and Marta Sabou	Austria
11:20-11:45	S06211 “Customizing Contextualized Language Models for Legal Document Reviews”	Shohreh shaghaghian, Luna Feng, Borna Jafarpour, and Nick Pogrebnyakov	Demark
11:45-12:00	S06216 “Clustering-based Automatic Construction of Legal Entity Knowledge Base from Contracts”	Fuqi Song and Éric de la Clergerie	France
12:00-12:15	S06215 “LegalOps: A Summarization Corpus of Legal Opinions”	Andrew Gargett, Robert Firth, and Nikolaos Aletras	UK
12:15-12:30	S06208 “AI Legal Counsel to train and regulate legally constrained autonomous systems”	Hal Ashton	UK
12:30-13:30	Noon Break		
13:30-13:55	Invited Speakers		
13:55-14:20	Application of Predictive Coding in Document Review Across Related Legal Matters - A Case Study	Robert Neary	USA
14:20-14:45	S06217 “Responsible AI: A Primer for the Legal Community”	Ilana Golbin, Anand Rao, Ali Hadjarian, and Daniel Kritzman	USA
14:45-15:10	S06207 “AI in the Law: Towards Assessing Ethical Risks”	Steven Wright	USA
14:10-15:35	S06218 “Application of Deep Learning in Recognizing Bates Numbers and Confidentiality Stamping from Images”	Christian J. Mahoney, Katie Jensen, Fusheng Wei, Haozhen Zhao, Han Qin, and Shi Ye	USA
15:35-15:55	S06213 “A bibliometric network analysis of Deep Learning publications applied into legal documents”	Alfredo Montelongo and João Luiz Becker	Brazil
15:55-16:00	Closing Remarks		

The 4th International Workshop on Big Data Analytic for Cyber Crime Investigation and Prevention <i>Thursday, December 10th, 2020; Location: online (USA Eastern Standard time zone)</i> <i>Workshop Chairs: Andrii Shalaginov, Asif Iqbal, Igor Kotsiuba, Mamoun Alazab</i> <i>Norwegian University of Science and Technology, Gjøvik, Norway</i> <i>KTH Royal Institute of Technology in Stockholm, Sweden</i> <i>Pukhov Institute for Modelling in Energy Engineering, Ukraine</i> <i>Charles Darwin University, Australia</i>		
Time (EST)	Title of the talk	Presenter/Author
09:00-09:05	Opening Remarks and Welcome	
		Andrii Shalaginov
09:05-09:15	S08216: Modern Cybercrime Investigation: technological advancement of Smart Devices and legal aspects of corresponding digital transformation	Andrii Shalaginov, Marina Shalaginova, Aleksandar Jevremovic, Marko Krstic
09:15-09:30	S08203: Towards verifiable evidence generation in forensic-ready systems	Lukas Daubner, Martin Macak, Barbora Buhnova, and Tomas Pitner

09:30-09:55	S08205: Intensive Image Malware Analysis and Least Significant Bit Matching Steganalysis	Yogesh Kulkarni and Anurag Gorkar
09:55-10:20	S08214: Detection of Previously Unseen Malware using Memory Access Patterns Recorded Before the Entry Point	Sergii Banin and Geir Olav Dyrkolbotn
10:20-10:45	S08204: Cyber Espionage: Reloaded	Keynote speaker: Ashwin Ram, Check Point Software Technologies, Australia
10:45-11:10	S08213: Real-Time Detection of Fake-Shops through Machine Learning	Louise Beltzung, Andrew Lindley, Olivia Dinica, Nadin Hermann, and Raphaela Lindner
11:10-11:30	Coffee Break	
11:30-11:55	S08215: CTI-Twitter: Gathering Cyber Threat Intelligence from Twitter using Integrated Supervised and Unsupervised Learning	Linn-Mari Kristiansen, Vinti Agarwal, Katrin Franke, and Raj Sanjay Shah
11:55-12:10	S08211: Smart Home Forensics: An Exploratory Study on Smart Plug Forensic Analysis	Asif Iqbal, Johannes Olegård, Ranjana Ghimire, Shirin Jamshir, and Andrii Shalaginov
12:10-12:35	S08212: Never-Ending Learner for Malware Analysis (NELMA)	Moumita Kamal and Douglas Talbert
12:35-13:00	Live: Using AI to protect children from online threats	Keynote speaker: Aleksandar Jevremovic, Singidunum University, Serbia
13:00-13:45	Lunch break	
13:45-14:10	S08210: Centrality and Scalability Analysis on Distributed Graph of Large-Scale E-mail Dataset for Digital Forensics	Selim Ozcan, Merve Astekin, Narasimha Karpoor Shashidhar, and Bing Zhou
14:10-14:35	S08209: Exploring Open Source Information for Cyber Threat Intelligence	Victor Adewopo, Bilal Gonen, and Festus Adewopo
14:35-15:00	S08202: Machine Learning Methods for Anomaly Detection in Industrial Control Systems	Johnathan Tai, Izzat Alsmadi, Yunpeng Zhang, and Fengxiang Qiao
15:00-15:20	Coffee Break	
15:20-15:35	S08201: National Security Intelligence through Social Network Data Mining	Nnaemeka Ekwunife
15:35-15:50	S08206: Financial Fraud Detection using Deep Support Vector Data Description	Masoud Erfani, Farzaneh Shoeleh, and Ali A. Ghorbani
15:50-15:55	Closing Remarks	Andrii Shalaginov

Workshop on Smart Farming, Precision Agriculture, and Supply Chain (SmartFarm-2020) <i>Workshop Chairs: Dr. Sudip Mittal, Dr. Maanak Gupta, Dr. Mahmoud Abdelsalam</i>		
Time (in EST)	Title	Presenter/Author
1:30pm – 1:35pm	Welcome to SmartFarm 2020	
1:35pm - 2:00pm	Automated Corn Ear Height Prediction Using Video-Based Deep Learning (S09201)	Hao Sha, Johnson Wong, Mohammad Al Hasan, George Mohler, Steve Becker, and Curtis Wiltse
2:00pm - 2:25pm	An exploratory analysis on Agritech policies, innovations and funding for climate change mitigation (S09202)	Pedro V Hernández Serrano, Lina Altenburg, and Parveen Kumar
2:25pm – 2:50pm	YieldPredict: A Crop Yield Prediction Framework for Smart Farms (S09203)	Nitu Kedarmal Choudhary, Sai Sree Laya Chukkapalli, Sudip Mittal, Maanak Gupta, Mahmoud Abdelsalam, and Anupam Joshi
2:50pm – 3:15pm	Development of A Wireless Power Transmission System for Agriculture Sensor Devices (S09204)	Charles Robinson, Brandon Nieman, Robert Craven, Muhammad Enagi Bima, and Charles Van Neste
3:15pm – 3:40pm	A Game Theoretic Analysis for Cooperative Smart Farming (S09205)	Deepti Gupta, Paras Bhatt, and Smriti Bhatt
3:40pm – 4:05pm	Automating Tiny ML Intelligent Dairy Sensors DevOPS Using Microsoft Azure (BigD396 ->S09)	Chandrasekar Vuppapapati, Anitha Ilapakurti, Karthik Chillara, Vanaja Mamidi, and Sharat Kedari
4:15 pm	Keynote - Opportunities and Challenges Developing Smart Farm Systems for Small-scale Agriculture Production	Michael P. Nattrass
	Closing Remarks	

The 2nd International Workshop on Big Data Tools, Methods, and Use Cases for Innovative Scientific Discovery (BTSD) 2020 <i>Sangkeun Lee, Travis Johnston: 12/10/2020 9am – 2:30 Presentation (15 minutes) & Q/A (5 minutes)</i>		
Time	Title	Presenter/Author
9:00am -9:20am	Opening Remarks	Matt Lee / Travis Johnston
9:20am – 9:40am	Pravallika Devineni, Bill Kay, Hao Lu, Anika Tabassum, Supriya Chintavali, and Sangkeun Lee, Toward Quantifying Vulnerabilities in Critical Infrastructure Systems (Paper ID: S18206)	Pravallika Devineni
9:40am –10:00am	Ohyung Kwon, Nayeon Lee, and Kangil Kim, Effect of pre-training to build a regression model using shallow neural network for semiconductor plasma etch process equipment (Paper ID: S18207)	Ohyung Kwon
10:00am –10:20am	Richard Valente, Alissa Ostapenko, Bryer Sousa, Jack Grubbs, Christopher Massar, Danielle Cote, and Rodica Neamtu, Classifying Powder Flowability for Cold Spray Additive Manufacturing Using Machine Learning (Paper ID: BigD666)	Richard Valente
10:20am—10:30am	Coffee Break	
10:30am –10:50am	Alvin Alaphat and Meng Jiang, SmartFund: Predicting Research Outcomes with Machine Learning and Natural Language Processing (Paper ID: S18201)	Meng Jiang
10:50am—11:10am	Biao Yin, Thomas Considine, Fatemeh Emdad, John Kelley, Robert Jensen, and Elke Rundensteiner, Corrosion Assessment: Data Mining for Quantifying Associations between Indoor Accelerated and Outdoor Natural Tests (Paper ID: S18205)	Biao Yin
11:10am—11:30am	António Correia, Shoaib Jameel, Daniel Schneider, Hugo Paredes, and Benjamim Fonseca, A Workflow-Based Methodological Framework for Hybrid Human-AI Enabled Scientometrics (Paper ID: S18208)	António Correia
11:30am—11:50am	Mohit Kumar, Lee Sangkeun, Park Byung, James Blum, Merry Ward, and Jonathan R. Nebeker, Advanced Health Information Technology Analytic Framework and Application to Hazard Detection (Paper ID: S18212)	Mohit Kumar
11:50am—1:00pm	Lunch Break	
1:00pm—1:20pm	David Tatis, Heidy Sierra, and Emmanuel Arzuaga, Residual Neural Network Architectures to Improve Prediction Accuracy of Properties of Materials (Paper ID: S18216)	Emmanuel Arzuaga
1:20pm—1:40pm	Latreche Othmane and Boukraa Doulkifli, Self-Service, On-Demand Creation of OLAP Cubes over Big Data: a Metadata-Driven Approach (Paper ID: S18217)	Latreche Othmane
1:40pm—2:00pm	Stefano Di Frischia, Paolo Giammatteo, Federico Angelini, Valeria Spizzichino, Elena De Santis, and Luigi Pomante, Enhanced Data Augmentation using GANs for Raman Spectra Classification (Paper ID: S18202)	Stefano Di Frischia
2:00pm – 2:20pm	Brian Cho, Teresa Dayrit, Yuan Gao, Zhe Wang, Tianzhen Hong, Alex Sim, and Kesheng Wu, Effective Missing Value Imputation Methods for Building Monitoring Data (Paper ID: BigD675)	Brian Cho
2:20pm—2:30pm	Closing Remarks	

Workshop on IoT based Big Data Architectures and Applications

Workshop Chairs: Hasan Ali Khattak <hasan.alikhattak@seecs.edu.pk>

Time	Title	Presenter/Author
9:00	S26202: Big Data Architectures for Vehicle Data Analysis	Christian Prehofer <c.prehofer@eu.denso.com>
9:30	S26204: Edge Computing with Big Data Cloud Architecture: A Case Study in Smart Building	Catherine Inibhunu <catherine.inibhunu@ontariotechu.ca>
10:00	S26205: Developing an Architecture for IoT Interoperability in Healthcare: A Case Study of Real-time SpO2 Signal Monitoring and Analysis	Quoc Nguyen <quoc.nguyen@ndsu.edu>
	Closing Remarks	

Applications of Big Data Technology in the Transport Industry

Workshop Chair: John Easton, University of Birmingham, UK

Time	Title	Presenter/Author
09:00 – 09:20	N235 - The Impact of COVID-19 on Flight Networks	Hiroki Kanezashi
09:20 – 09:40	S10202 - High Beam Detection Method Based on Deep Learning	Liu Peishun
09:40 – 10:00	S10204 - Towards Modular Modelling and Analytic for Multi-Modal Transportation Networks	Ramyar Saeedi
10:00 – 10:10	Q&A Papers 1 - 3	
10:10 – 10:15	Comfort Break	
10:15 – 10:35	S10201 - Metadata-Driven Industrial-Grade ETL System	Peter Panfilov
10:35 – 10:55	S10203 - SELIS BDA: Big Data Analytics for the Logistics Domain	Nikodimos Provatas
10:55 – 11:15	S10205 - Blockchain Application in Remote Condition Monitoring	Rahma Alzahrani
11:15 – 11:25	Q&A Papers 4 - 6	
11:25 – 11:50	Coffee Break	
11:50 – 12:10	BigD564 - An Attack Detection Mechanism for Intelligent Transport System	Edivaldo Valentini
12:10 – 12:30	BigD461 - Algorithm-centered Approach to Improve Track Performance Monitoring with Rail Profile Data	Ahmed Lasisi
12:30 – 12:50	S10206 - From Spatial-Temporal Cluster Relationships to Lifecycles: Framework and Mobility Applications	Ivens Portugal
12:50 – 13:10	S10207 - A Quantum Bayesian Approach for Bike Sharing Demand Prediction	Ramkumar Harikrishnakumar
13:10 – 13:25	Q&A Papers 7 - 10	
13:25 – 13:30	Closing Remarks	

GTA³ 4.0: The 4th workshop on Graph Techniques for Adversarial Activity Analytics

Workshop Chairs: Jiejun Xu, Hanghang Tong, Andrea Bertozzi, Vince Lyzinski, George Chin, Joel Douglas

Time (US EST)	Title	Presenter/Author
12:15 pm	Opening Remarks	Workshop Organizers
12:20 pm	Keynote 1: TBD	Prof. Leman Akoglu
1:00 pm	Keynote 2: TBD	Dr. Patrick Rubin-Delanchy
1:40 pm	Inexact Attributed Subgraph Matching (S12211)	Thomas Tu, Jacob Moorman, Dominic Yang, Qinyi Chen, and Andrea Bertozzi
1:55 pm	Robust and Scalable Entity Alignment in Big Data (S12202)	James Flamino, Christopher Abriola, Benjamin Zimmerman, Zhongheng Li, and Joel Douglas
2:10 pm	Multi-Channel Entity Alignment via Name Uniqueness Estimation (S12201)	Miquette Orren, Patrick Mackey, Natalie Heller, and George Chin
2:20	Evaluation of Alignment: Precision, Recall, Weighting and Limitations (S12204)	Joseph Cottam, Natalie Heller, Christopher Ebsch, Rahul Deshmukh, Patrick Mackey, and George Chin

2:30	Using Graph Edit Distance for Noisy Subgraph Matching of Semantic Property Graphs (S12205)	Christopher Ebsch, Joseph Cottam, Natalie Heller, Rahul Deshmukh, and George Chin
2:40	Coffee Break	
3:00	Keynote 3: Graph Representation Learning: Recent Advances and Open Challenges	Prof. William Hamilton
3:40	Keynote 4: Exploring Rare Categories on Graphs: Local vs. Global	Prof. Jingrui He
4:20 pm	Graph Adversarial Attacks and Defense: An Empirical Study on Citation Graph (S12210)	Chau Pham, Vung Pham, and Tommy Dang
4:35 pm	Data-Driven Template Discovery Using Graph Convolutional Neural Networks (S12206)	Mikel Joaristi, Sumit Purohit, Rahul Deshmukh, and George Chin
4:50 pm	Semantic Guided Filtering Strategy for Best-effort Subgraph Matching in Knowledge Graphs (S12207)	Alexei Kopylov, Jiejun Xu, Kangyu Ni, Shane Roach, and Tsai-Ching Lu
5:00 pm	Fault-Tolerant Subgraph Matching on Aligned Networks (S12208)	Thomas Tu and Dominic Yang
5:10 pm	Static and Dynamic Social Network Models for the Analysis of Transshipment in Illegal Fishing (S12209)	Stefano Stamato and Andrew Park
5:20 pm	Who killed Lilly Kane? A case study in applying knowledge graphs to crime fiction. (S12203)	Mariam Alaverdian, William Gilroy, Veronica Kirgios, Xia Li, Carolina Matuk, Daniel Mckenzie, Tachin Ruangkriengsin, Andrea Bertozzi, and Jeffrey Brantingham
5:30 pm	Closing Remarks	

3rd Workshop on Energy-Efficient Machine Learning and Big Data Analytics		
<i>Workshop Chair: Mohammed Alawad</i>		
Time	Title	Presenter/Author
1:30-2:30PM	Keynote Speech Accurate, Real-time Energy-efficient Scene Perception through Hardware Acceleration	Iris Bahar
2:30-2:55	(S11201) SWVBiL-CRF: Selectable Word Vectors-based BiLSTM-CRF Power Defect Text Named Entity Recognition	JianBin Li, SuWan Fang, YuQi Ren, KunChang Li, and MingYu Sun
2:55-3:20	(S11202) Energy-Efficient Energy Analytics Using a General Purpose Graphics Processing Unit	Sagnik De and Wojciech Golab
3:20-3:45	(BigD235) Increasing Prediction Accuracy for Human Activity Recognition Using Optimized Hyperparameters	Niyati R. Darji and Samuel A. Ajila
3:45-4:10	(BigD608) Towards Green Query Processing - Auditing Power Before Deploying	Simon Pierre Dembele, Ladjel Bellatreche, and Carlos Ordonez
	Closing Remarks	

8th International Workshop on Distributed Storage and Blockchain Technologies for Big Data		
<i>Workshop Chairs: Hui Li</i>		
Time	Title	Presenter/Author
Dec 11 9:00 – 9:20 am	S13202: A Blockchain-Based Approach for Dynamically Differential Authorized Deduplication	Tian Zhao, Hui Li, Xin Yang, Han Wang, and Kaixuan Xing

Dec 11 9:20 – 9:40 am	S13201: A Dynamic Sharding Protocol Design for Consortium Blockchains	Zhixuan Zhou, Zhijie Qiu, Qiang Yu, and Hong Chen
Dec 11 9:40 – 10:00 am	S13205: Implementation of a Multi-tenancy Adaptive Cache Admission Model	Huajun Ma, Jiyang Zhang, and Zixian Wang
Dec 11 10:00 – 10:20 am	Coffee Break	
Dec 11 10:20 – 10:40 am	S13206: Dynamic Load Balance for Hot-spot and Unbalance Region Problems in HBase	Yibo Chen, Xing Xiang, Xiao Ling, Xiangchi Zhang, Fan Wu, and Jianliang Gao
Dec 11 10:40 – 11:00 am	S13207: On Chord in Distributed Storage System	Haiyang Yu, Hui Li, Xin Yang, Huajun Ma, Bin Wang, and Jia Chen
Dec 11 11:00 – 11:20 am	Keynote speech	
	Closing Remarks	

3rd Workshop on Big Data for CyberSecurity (BigCyber 2020) <i>Workshop Chairs: Dr. Karuna Joshi and Dr. Sudip Mittal</i>		
Time (in EST)	Title	Presenter/Author
9:00am – 9:10am	Welcome to BigCyber 2020	
9:10am – 9:35am	BEAM: An Anomaly-Based Threat Detection System for Enterprise Multi-Domain Data (S14201)	Derek Lin, Anying Li, and Ryan Foltz
9:35am – 10:00am	An Approach to Developing EEG-Based Person Authentication System (S14203)	Meetkumar Patel and Mohammad Husain
10:00am - 10:25am	A Comparative Study of Deep-Learning based Named Entity Recognition Algorithms for Cybersecurity (S14204)	Soham Dasgupta, Aritran Piplai, Anantaa Kotal, and Anupam Joshi
10:25am -10:50am	Using Knowledge Graphs and Reinforcement Learning for Malware Analysis (S14205)	Aritran Piplai, Priyanka Ranade, Anantaa Kotal, Sudip Mittal, Sandeep Narayanan, and Anupam Joshi
10:50am-11:15am	Identifying Influences on Information Security Framework Adoption: Applying a Modified UTAUT Model (BigD428 -> S14)	William Lidster and Shawon Rahman
11:30am	<i>Keynote</i> - Data Science, COVID-19 Pandemic, Privacy and Civil Liberties (S14202)	Bhavani Thuraisingham
After the keynote	Closing Remarks	

International Workshop on Big Data Reduction <i>Workshop Chairs: Dingwen Tao, Sheng Di</i>		
Time	Title	Presenter/Author
09:00 – 09:10 am	Opening Remarks and Welcome	Dingwen Tao, Sheng Di
09:10 – 09:55 am	Keynote Speech Title: TBD	Bill Spotz

		Program Manager, Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy
09:55 – 10:20 am	S15203: SDRBench: Scientific Data Reduction Benchmark for Lossy Compressors	Kai Zhao, Sheng Di, Xin Liang, Sihuan Li, Dingwen Tao, Julie Bessac, Zizhong Chen, and Franck Cappello
10:20 – 10:45 am	S15204: Assessing Differences in Large Spatio-temporal Climate Datasets with a New Python package	Alexander Pinard, Dorit Hammerling, and Allison Baker
10:45 – 11:10 am	BigD467: Accelerating Text Mining Using a Domain-Specific Stop Word List	Farah Alshanik, Amy Apon, Alexander Herzog, Ilya Safro, and Justin Sybrandt
11:10 – 11:35 am	S15205: LCFI: A Fault Injection Tool for Studying Lossy Compression Error Propagation in HPC Programs	Baodi Shan, Aabid Shamji, Jiannan Tian, Guanpeng Li, and Dingwen Tao
11:35 – 11:50 am	S15206: A Middleware Approach to Leverage Distributed Data Deduplication Capability on HPC and Cloud Storage Systems	Hsing-bung Chen, Sihai Tang, and Song Fu
11:50 – 12:30 pm	Lunch Break	
12:30 – 12:55 pm	BigD700: Interpretable Visualization and Higher-Order Dimension Reduction for ECoG Data	Kelly Geyer, Frederick Campbell, Andersen Chang, John Magnotti, Michael Beauchamp, and Genevera Allen
12:55 – 13:20 pm	S15207: Efficient Data Management in Neutron Scattering Data Reduction Workflows at ORNL	William Godoy, Peter Peterson, Steven Hahn, and Jay Billings
13:20 – 13:45 pm	S15210: Topology Preserving Data Reduction for Computing Persistent Homology	Nicholas Malott, Aaron Sens, and Philip Wilsey
13:45 – 14:10 pm	S15211: Combining Spatial and Temporal Properties for Improvements in Data Reduction	Megan Hickman Fulp, Ayan Biswas, and Jon Calhoun
14:10 – 14:35 pm	S15213: A Fast, Scalable, Universal Approach for Distributed Data Aggregations	Niranda Perera, Vibhatha Abeykoon, Chathura Widanage, Supun Kamburugamuve, Thejaka Amila Kanewala, Pulasthi Wickramasinghe, Ahmet Uyar, Hasara Maithree, Damitha Lenadora, and Geoffrey Fox
14:35 – 14:40 pm	Closing Remarks	

The Fourth IEEE International Workshop on Benchmarking, Performance Tuning and Optimization for Big Data Applications (BPOD), Thursday, December 10 <i>Workshop Chairs: Zhiyuan Chen, Jianwu Wang, Feng Chen, Yiming Ying</i> Workshop website: https://userpages.umbc.edu/~jianwu/BPOD		
Time (US Eastern Time)	Title	Presenter/Author
	Session 1 (Analytics & Optimization)	
9-9:05 am	Opening Remark	Workshop Chairs
9:05-9:25 am	S17212: Autoencoder-based outlier detection for sparse, high dimensional data	Wanghu Chen, Huijun Li, Jing Li, Ali Arshad
9:25-9:45 am	S17206: Impact of Map-Reduce framework on Hadoop and Spark MR Application Performance	Ishaan Lagwankar, Ananth Narayan Sankaranarayanan, Subramaniam Kalambur
9:45-10:00 am	S17207: A Topology-Aware Performance Prediction Model for Distributed Deep Learning on GPU Clusters	Zheyu Lin, Xukun Chen, Hanyu Zhao, Yunteng Luan, Zhi Yang, Yafei Dai,
10:00-10:15am	S17208: Benchmarking performance of RaySGD and Horovod for big data applications	Shruti Kunde, Amey Pandit, Rekha Singhal
	Break	
	Session 2 (Analytics & Optimization)	
10:30-10:50	S17209: Not Half Bad: Exploring Half-Precision in Graph Convolutional Neural Networks	John Brennan, Stephen Bonner, Amir Atapour-Abarghouei, Philip Jackson, Boguslaw Obara, Andrew Stephen McGough
10:50-11:10	S17203: Exploring Characteristics of Inter-cluster Machines and Cloud Applications on Google Clusters	Yuhui Lin, Adam Barker, Sheriffo Ceesay

11:10-11:30	BigD601: Making an Array Database Language Server-Side Extensible	Otoniel José Campos Escobar, Peter Baumann, Dimitar Misev
11:30-11:45	S17213: Towards Collaborative Optimization of Cluster Configurations for Distributed Dataflow Jobs	Jonathan Will, Jonathan Bader, Lauritz Thamsen
	Lunch Break	
1-2 pm	S17216: Keynote	Geoffrey Fox, Indiana University Bloomington
	Session 3 (Benchmark)	
2:00-2:20 pm	BigD595: Evaluating NoSQL Systems for Decision Support: An Experimental Approach	Tomas Llano-Rios, Mohamed Khalefa, Antonio Badia
2:20-2:40 pm	S17204: Benchmarking HOAP for Scalable Document Data Management: A First Step	Yifan Tian, Michael Carey, Ian Maxon
2:40-2:55 pm	S17205: Performance Benchmarking of Automated Sentence Denoising using Deep Learning	Zhantong Liang, Abdou Youssef
2:55-3:10 pm	S17211: RDFINT: A Benchmark for Comparing Data Warehouse with Virtual Integration Approaches for Integration of RDF Data	Samson Oni, Kajal Pansare, Sukrit Arneja, Zhiyuan Chen, Adina Crainiceanu, Don Needham
3:10-3:25 pm	Break	
	Session 4 (Analytics & Optimization)	
3:25-3:45 pm	BigD541: Computational performance of heterogeneous ensemble frameworks on high-performance computing platforms	Linhua Wang, Prem Timsina, Gaurav Pandey
3:45-4:05 pm	N246: Improving Join Reordering for Large Scale Distributed Computing	Amogh Margoor, Mayur Bhosale
4:05-4:20 pm	S17214: A Simple Low Cost Parallel Architecture for Big Data Analytics	Carlos Ordonez, Sikder Tahsin Al-Amin, and Xiantian Zhou
4:20-4:35 pm	S17210: Uncovering Performance Interference of Multi-Tenants in Big Data Environments	Adriano Lange, Tiago Rodrigo Kepe, Eduardo Cunha de Almeida, Marcos Sfair Sunye
4:35-4:50 pm	S17201: Automatic Tuning of Hyperparameters for Neural Networks in Serverless Cloud	Alex Kaplunovich, Yelena Yesha
	Closing Remarks	

The Fourth IEEE Workshop on Human-in-the-loop Methods and Future of Work in BigData (HMData 2020) <i>Workshop Chair: Senjuti Basu Roy (New Jersey Institute of Technology), Alex Quinn (Purdue University), Atsuyuki Morishima (University of Tsukuba)</i> <i>The latest program is available at https://humanmachinedata.org</i>		
Time	Title	Author
1:00PM	Opening	
1:05	Keynote: Crowd Sleuths: Solving Mysteries with Crowdsourcing, Experts, and AI	Kurt Luther (Virginia Tech)
1:55	Short Break	
2:00	[S20213] Human-in-the-Loop Business Modelling for Emergent External Factors	Shubhi Asthana, Shikhar Kwatra, Christine T. Wolf, Pawan Chowdhary, and Taiga Nakamura
2:20	[S20215] Validation of CyborgCrowd Implementation Possibility for Situation Awareness in Urgent Disaster Response -Case Study of International Disaster Response in 2019	Munenari Inoguchi, Keiko Tamura, Kousuke Uo, and Masaki Kobayashi

2:40	Break	
3:00	[S20217] Interpretation of Sentiment Analysis with Human-in-the-Loop	Vijaya Kumari Yeruva, Mayanka Chandrashekar, Yugyung Lee, Jeff Rydberg-Cox, Virginia Blanton, and Nathan A Oyler
3:20	[S20211] Exploring Contextual Paradigms in Context-Aware Recommendations	Conor Morgan, Iulia Paun, and Nikos Ntarmos
3:40	[S20208] Analysis of Hand-drawn Maps of Places in Natural Disaster Pictures	Seungun Kim, Masaki Matsubara, and Atsuyuki Morishima
3:52	[S20209] Human-in-the-loop Approach towards Dual Process AI Decisions	Hikaru Uchida, Masaki Matsubara, Kei Wakabayashi, and Atsuyuki Morishima
4:04	[S20205] Modelling Hybrid Human-Artificial Intelligence Cooperation: A Call Center Customer Service Case Study	Laura H. Kahn, Onur Savas, Adamma Morrison, Kelsey A. Shaffer, and Lila Zapata
4:16	Break	
4:40	[S20207] General Self-aware Information Extraction from Labels of Biological Collections	Icaro Alzuru, Andréa Matsunaga, Maurício Tsugawa, and José A.B. Fortes
5:00	[BigD439] Predicting Skill Shortages in Labor Markets: A Machine Learning Approach	Nikolas Dawson, Marian-Andrei Rizoiu, Benjamin Johnston, and Mary-Anne Williams
5:20	[S20203] Adaptive Image Scaling for Corresponding Points Matching between Images with Differing Spatial Resolutions	Hisatoshi Toriya, Ashraf Dewan, and Itaru Kitahara
5:32	[S20206] A Framework for Recommendation Algorithms Using Knowledge Graph and Random Walk Methods	Takafumi Suzuki, Satoshi Oyama, and Masahito Kurihara
5:44	Closing Remarks	

The 5th IEEE International Workshop on Big Spatial Data (BSD 2020)		
<i>Workshop Chairs: Farnoush Banaei-Kashani, Siyuan Lu, Chengyang Zhang, Abdeltawab Hendawi</i>		
Time	Title	Presenter/Author
9:00-9:05	Opening	Workshop Organizers
9:05-9:25	Ad Recommendation utilizing user behavior in the physical space to represent their latent interest (Paper ID: S21205)	Takanobu Omura , Kenta Suzuki, Panote Siriaraya, Mohit Mittal, Yukiko Kawai, and Shinsuke Nakajima
9:25-9:55	Crowd Forecasting at Venues with Microblog Posts Referring to Future Events (Paper ID: S21206)	Ryotaro Tsukada , Haosen Zhan, Shonosuke Ishiwatari, Masashi Toyoda, Kazutoshi Umemoto, Haichuan Shang, and Koji Zettsu
9:55-10:25	Toward Identifying the Urban Community Structure From Population Flow and Public Services Distribution (Paper ID: BigD551)	Qinghe Liu , Zhicheng Liu, Yinfei Xu, Weiting Xiong, Junyan Yang, and Qiao Wang
10:25-10:45	Coffee Break	
10:45-11:15	Country-wide mobility changes observed using mobile phone data during COVID-19 pandemic (Paper ID: BigD271)	Georg Heiler , Tobias Reisch, Jan Hurt, Mohammad Forghani, Aida Omani, Allan Hanbury, and Farid Karimipour

11:15-11:45	Real-time Traffic Jam Detection and Congestion Reduction Using Streaming Graph Analytics (Paper ID: BigD712)	Zainab Abbas , Paolo Sottovia, Mohamad Al Hajj Hassan, Daniele Foroni, and Stefano Bortoli
11:45-12:05	Leveraging Differential Privacy in Geospatial Analyses of Standardized Healthcare Data (Paper ID: S21204)	Daniel Harris
12:05-12:25	Coffee Break	
12:25-13:25	Keynote (Paper ID: S21207)	Xin Chen
13:25-13:30	Closing Remarks	

The 4th IEEE International Workshop on Big Data for Financial News and Data

Date: Thursday, Dec. 10, 2020

Workshop Chairs: Quanzhi Li, Xiaozhong Liu, Sameena Shah

Time	Title	Presenter/Author
9:00AM - 9:20AM	S43207 - A Novel Hybrid Approach with A Decomposition Method and The RVFL Model for Crude Oil Price Prediction	Chengyuan Zhang, Fuxin Jiang, Shouyang Wang, Wei Shang
9:20AM - 9:45AM	S43201 - DeepCatch: Predicting Return Defaulters in Taxation System using Example-Dependent Cost-Sensitive Deep Neural Networks	Priya Mehta, Ch Sobhan Babu, S.V. Kasi Visweswara Rao, Sandeep Kumar
9:45AM - 10:10AM	S43208 - Hardening Soft Information: A Transformer-Based Approach to Forecasting Stock Return Volatility	Matthew Caron and Oliver Müller
10:10AM -10:35AM	BigD375 - Exploring market power using deep reinforcement learning for intelligent bidding strategies	Alexander Kell, Matthew Forshaw, and A. Stephen McGough
10:35AM -11:00AM	BigD582 - Predicting Stock Price Movements with Text Data using Labeling based on Financial Theory	Fredrik Ahnve, Kasper Fantenberg, Gustav Svensson, Daniel Hardt
11:00AM-11:15AM	Coffee Break	
11:15AM-11:40AM	BigD614 - Repeating Link Prediction over Dynamic Graphs	Daniele Montesi, Vladimir Vlassov, and Sarunas Girdzijauskas
11:40AM-12:05PM	S43205 - Anomaly Detection in Exchange Traded Funds	Nitish Bahadur, Randy Paffenroth
12:05PM-12:30PM	BigD532 - Large Scale Financial Filing Analysis on HPC Systems	Matthias Murray, Lili Xu, Arjuna Chala, and Roger Dev
12:30PM-12:55PM	S43206 - Causal Maps for Multi-Document Summarization	Sasha Strelnikoff, Aruna Jammalamadaka, Dana Warmusley
12:55PM-13:20PM	S43209 - Large-scale Data-driven Segmentation of Banking Customers	Md Monir Hossain, Mark Sebestyen, Dhruv Mayank, Omid Ardakanian, Hamzeh Khazaei
	Closing Remarks	

IEEE Big Data Workshop on Data Science for Smart and Connected Communities

December 10, 2020

Time	Title	Presenter/Author
1:30-1:40pm	Opening Remarks	George Mohler
	Session 1: Crime Modeling	
1:40:155pm	S23208 A Self-limiting Hawkes Process: Interpretation, Estimation, and Use in Crime Modeling	Jack Olinde
2pm-2:15pm	S23203 Repurposing recidivism models for forecasting police officer use of force	Samira Khorshidi
2:20pm-2:35pm	S23204 Analyzing Effectiveness of Gang Interventions using Koopman Operator Theory	Sian Wen
2:40pm-2:55pm	S23202 Interpretable Hawkes Process Spatial Crime Forecasting with TV-Regularization	Hao Sha
3pm-3:15pm	S23205 Emotion Classification and Textual Clustering Techniques for Gang Intervention Data	Ruofei Wu

3:15pm-3:30pm	Coffee Break	
3:30pm-4:15pm	Invited Panel: Shannon Reid, Jeremy Carter, and Jeff Brantingham	
4:15pm-4:30pm	Coffee Break	
	Session 2: Covid-19 and Mobility Modeling	
4:30pm-4:45pm	S23207 <i>Contact-Tracing based on Time-Varying Graphs Analysis</i>	Lorenzo Goglia
4:50pm-5:05pm	BigD329 <i>3D-CLoST: A CNN-LSTM Approach for Mobility Dynamics Prediction in Smart Cities</i>	Stefano Fiorini
5:10pm-5:25pm	S23201 <i>An Investigation of Containment Measures Against the COVID-19 Pandemic in Mainland China</i>	Ji LIU
5:30pm-5:45pm	S23206 <i>Analysis and Classification of Vaccine Dialogue in the Coronavirus Era</i>	Nijhum Paul
5:50pm-6pm	Closing Remarks	

PEASH'20 Workshop (Dec 11 th 2020)			
Chairs: Hui Zhang, Hongfeng Yu, Weijia Xu			
Time	PaperID	Title	Presenter/Author
9:00-9:25	S24201	Service Provisioning through High Level, Complexity Hiding Interfaces	Guangchen Ruan, et al.
9:25-9:50	S24207	Cosine similarity distance pruning algorithm Based on graph attention mechanism	Huaxiong Yao, et al.
9:50-10:15	S24203	Facilitating the HPC Data Center Host Efficiency through Big Data Analytics	Jack Rager, et al.
10:15-10:40	S24202	A Sentiment Analysis Service Platform for Streamed Multilingual Tweets	Ioanna Karageorgou, et al.
10:40-11:00		Coffee Break	
11:00-11:25	S24204	Web-Based Interactive Visualization for Data Structure in Python	Juan Lin, et al..
11:25-11:50	S24206	Baseline for Performance Prediction of Android Applications	Anders Skretting, et al.
11:50-12:15	S24205	A Multivariate Uncertainty Visualization Platform for Exploring Nitrogen Leaching	Babak Samani, et al.
12:15-12:40	S24208	RIVACHain: Blockchain-based Integrity Verification for Scientific Data Transfers	Ahmed Alhussen, et al
		Closing Remarks	

The 3rd International Workshop on Big Media Dataset Construction, Management and Applications

<i>Workshop Chairs: Mingyu You, Ruijiang Li</i>		
Time	Title	Presenter/Author
13:30-13:50	S25201: Detecting Curve Text with Local Segmentation Network and Curve Connection	Zhao Zhou, Hao Ye, Luhui Chen, and Yingbin Zheng
13:50-14:10	S25202: FSRGAN-DB: Super-resolution Reconstruction Based on Facial Prior Knowledge	Wengang Zhou, Chaoqun Hong, Xiaodong Wang, and Zhiqiang Zeng
14:10-14:30	S25203: Development and Application of an Intensive Care Medical Data Set for Deep Learning	Shangping Zhao, Pan Liu, Guanxiu Tang, Yanming Guo, and Guohui Li
14:30-14:50	S25204: Multi-level Feature Fusion Network for Single Image Super-Resolution	Xinxia Zhang, Xiaoqin Zhang, Li Zhao, Runhua Jiang, Pengcheng Huang, and Jiawei Xu
14:50-15:10	S25205: Self-calibrated Attention Residual Network for Image Super-Resolution	Anqi Rong, Li Zhao, Pengcheng Huang, and Jiawei Xu
15:10-15:30	S25206: A Nonlocal Denoising Framework Based on Tensor Robust Principal Component Analysis with Lp norm	Mengqing Sun, Li Zhao, Jingjing Zheng, and Jiawei Xu
15:30-15:50	Coffee Break	
15:50-16:10	S25207: Feature Fusion Based on Sparse Block for Image Super-resolution	Shengping Wang, Li Zhao, Runhua Jiang, Pengcheng Huang, and Jiawei Xu
16:10-16:30	S25208: Look into Multi-Person: A New Benchmark for Pose Estimation and Human Parsing	Yanlu Cai, Yuqing Li, Runyu Peng, Yipei Xu, Chenzhe Jin, and Cheng Jin
16:30-16:50	S25209: Design of Fine-grained Plant Dataset and A Plant Image Acquisition Tool	Weilin Wan, Bingyu Tang, Ziheng Sun, and Haochen Ye
16:50-17:10	S25210: OSD: An Occlusion Skeleton Dataset for Action Recognition	Yuan Wu, Haoyue Qiu, Jing Wen, and Rui Feng
17:10-17:30	S25211: SOF: A Synthetic Occluded Face Dataset	Yuan Wu, Zi Yuan, Lurui Jin, and Rui Feng
17:30-17:35	Closing Remarks	

Computational Archival Science		
<i>Workshop Chairs: Mark Hedges, Victoria Lemieux, Richard Marciano</i>		
Friday 11 December		
Time	Title	Presenter/Author
9:00 – 9:15	Welcome	Mark Hedges (King's College London), Victoria Lemieux (U. British Columbia) Richard Marciano (U. Maryland),
9:15 – 10:00	Keynote	Dave de Roure
10:00 – 10:40	SESSION 1: Analytics for archival processing	
10:00 – 10:20	Automatic Extraction of Dublin Core Metadata from Presidential E-records (S02209)	W. Underwood
10:20 – 10:40	Multi-label Classification of Chinese Judicial Documents based on BERT (S02206)	M. Dai, C.L. Liu
10:40 – 11:00	Coffee Break (coffee not supplied!)	
11:00 – 11:40	SESSION 2: Analytics for archival processing (cont.)	
11:00 – 11:20	A Computational Approach to Historical Ontologies (S02208)	M. Kelly, J. Greenberg, C. B. Rauch, S. Grabus, J. P. Boone, J.A. Kunze, P. Melville Logan
11:20 – 11:40	Digital Curation and Machine Learning Experimentation in Archives (S02211)	T. Randby, R. Marciano
11:40 – 12:40	SESSION 3: Analysing historical data and documents	
11:40 – 12:00	Curation of Historical Arabic Handwritten Digit Datasets from Ottoman Population Registers: A Deep Transfer Learning Case Study (S02205)	Y. Said Can, M. Erdem Kabadayi
12:00 – 12:20	HRCenterNet: An Anchorless Approach to Chinese Character Segmentation in Historical Documents (S02212)	C-W Tang, C-L Liu, P-S Chiu
12:20 – 12:40	Towards Automatic Data Cleansing and Classification of Valid Historical Data: An Incremental Approach Based on MDD (BigD631)	E. O'Shea, R. Khan, C. Breathnach, T. Margaria

12:40 – 1:40	Lunch Break (lunch not supplied!)	
1:40 – 2:00	Discussion (and questions for keynote speaker)	
2:00 – 2:40	SESSION 4: Representation in archives	
2:00 – 2:20	Computational Treatments to Recover Erased Heritage: A Legacy of Slavery Case Study (S02203)	L. Perine, R. K. Gnanasekaran, P. Nicholas, A. Hill, R. Marciano
2:20 – 2:40	Elevating “Everyday” Voices and People in Archives through the Application of Graph Database Technology (S02207)	M. Conrad, L. Williams
2:40 – 3:20	SESSION 5: Visual and audio archives	
2:40 – 3:00	A Study of Spoken Audio Processing using Machine Learning for Libraries, Archives and Museums (S02202)	W. Xu, M. Esteva, P. Cui, E. Castillo, K. Wang, H-R Hopkins, T. Clement, A. Choate, R. Huang
3:00 – 3:20	From Computational De-Morphogenesis to Contaminated Representation for the Contemporary Digital Tectonics and Lexicon: Vulnerability and Resilience of the Generative Representation in the Web-Oriented Platform – BIM/GD (BigD606)	A. De Masi
3:20 – 3:40	Coffee Break (coffee not supplied!)	
3:40 – 4:20	SESSION 6: Web and social media archives	
3:40 – 4:00	Modeling Updates of Scholarly Webpages Using Archived Data (S02210)	Y. Jayawardana, A. C. Nwala, G. Jayawardena, J. Wu, S. Jayarathna, M.L. Nelson, C. Lee Giles
4:00 – 4:20	Using a Three-step Social Media Similarity (TSMS) Mapping Method to Analyze Controversial Speech Relating to COVID-19 in Twitter Collections (S02204)	Z. Yin, L. Fan, H. Yu, A. Gilliland
4:20 – 5:00	SESSION 7: Discussion	
5:00 – 5:20	Closing Remarks (Organizers)	

Big Food and Nutrition Data Management and Analysis (BFNDMA 2020)		
<i>Workshop Chairs: Tome Eftimov, Gorjan Popovski, Panče Panov, Barbara Koroušić Seljak</i>		
Time	Title	Presenter/Author
09:00 – 09:30	Invited talk: Exploring Food Contents in Scientific Literature with FoodMine	Giulia Menichetti
09:30 – 09:50	BuTTER: Bidirectional LSTM for Food Named-Entity Recognition (S29203)	Gjorgjina Cenikj
09:50 – 10:10	Discovering Entity Profiles Candidate for Entity Resolution on Linked Open Data Halal Food Products (S29202)	Nur Aini Rakhmawati and Ahmad Choirun Najib
10:10 – 10:30	Leveraging Social Network Analysis to Explore Obesity Talks on Twitter (S29201)	Edwin Mitei and Thanaa Ghanem
10:30 – 10:50	Coffee Break	
10:50 – 11:20	Invited talk: Semantic Annotation of Recipes to Investigate COVID-19 Impact on Food Consumption Process	Dragi Kocev
11:20 -11:50	Comparison of Feature Selection Algorithms for Minimization of Target Specific FFQs (S29207)	Nina Rešič
11:50 – 12:10	Exploring Knowledge Domain Bias on a Prediction Task for Food and Nutrition Data (S29206)	Gordana Ispirova
12:10 – 12:30	Coffee Break	
12:30 – 12:50	Toward Robust Food Ontology Mapping (S29204)	Gorjan Popovski

12:50 – 13:10	APRICOT: A humAn-comPuteR InterACtion tool for linking foOd wasTe streams across different (S29205)	Riste Stojanov
Closing Remarks		

International Workshop on Fair and Interpretable Learning Algorithms (FILA 2020) <i>Workshop Chairs: Arindam Pal, Yinglong Xia, Abhijnan Chakraborty and Mayank Singh</i>		
Time	Title	Presenter/Author
09:00-09:10	Welcome Talk	<i>Abhijnan and Mayank</i>
09:10-10:00	Keynote Talk: Network Derivative Mining	<i>Dr. Hanghang Tong</i>
10:00-10:25	S31212: Analyzing ‘Near Me’ Services: Potential for Exposure Bias in Location-based Retrieval	<i>Ashmi Banerjee, Gourab K Patro, Linus W. Dietz, and Abhijnan Chakraborty</i>
10:25-10:40	S31214: Interpretable Machine Learning for Understanding Epidemic Data	Dean Frederick Hougen, Jin-Song Pei, and Sai Teja Kanneganti
10:40-11:05	S31210: A Generic Framework for Black-box Explanations	<i>Clément Henin and Daniel Le Métayer,</i>
11:05-11:20	Coffee Break	
11:20-11:45	S31211: Fairness for Whom? Understanding the Reader's Perception of Fairness in Text Summarization	Anurag Shandilya, Abhisek Dash, Abhijnan Chakraborty, Kripabandhu Ghosh, and Saptarshi Ghosh,
11:45-12:00	S31205: Fairness Metrics: A Comparative Analysis	Pratyush Garg, John Villaseñor, and Virginia Foggo,
12:00-12:25	S31209: Validation of an Alternative Neural Decision Tree	Yu Leo Lu and Chunming Wang
12:25-12:50	S31206: BeFair: Addressing Fairness in the Banking Sector	Alessandro Castelnovo, Riccardo Crupi, Giulia Del Gamba, Greta Greco, Aisha Naseer, Daniele Regoli, and Beatriz San Miguel Gonzalez,
12:50-1:00	Closing Remarks	

2020 Workshop on Data Science in Medicine and Healthcare (DSMH 2020)		
<i>Workshop Chair: Xiong Liu</i>		
Time	Title	Presenter/Author
9-9:30	Keynote: The Big Short with AI in Biomedical Sciences: from Prediction to Action	Jiang Bian, Yi Guo, and Mattia Proserpi
9:30-9:50	Personal Digital Life Coach for Physical Therapy	Goce Popovski, Vasco Ponciano, Gonalo Marques, Ivan Miguel Pires, Eftim Zdravevski, and Nuno M. Garcia
9:50-10:10	Control and Prevention of Personal Stress	Hugo Marques, Hugo Carvalho, Jos� Morgado, Nuno M. Garcia, Ivan Miguel Pires, and Eftim Zdravevski
10:10-10:30	Features Importance to Improve Interpretability of Chronic Kidney Disease Early Diagnosis	Pedro A. Moreno-Sanchez
10:30-10:50	Comparative analysis of machine learning methods for analyzing security practice in electronic health records' logs	Prosper K. Yeng, Muhammad Ali Fauzi, and Bian Yang
10:50-11:10	E-health and M-health applications in Georgia: A review on the free available applications for Android Devices	Salome Oniani, Gonalo Marques, Ivan Miguel Pires, Salome Muhkashavria, and Nuno M. Garcia
11:10-11:30	Diseases identification with big data concept – The older people community	Andr� Esteves, Vasco Ponciano, and Ivan Miguel Pires
11:30-11:50	Coffee Break	
11:50-12:10	Exploration of links between anxiety purchases, deprivation and personality traits	Vanja Ljevar, James Goulding, and Gavin Smith
12:10-12:30	Reducing Reporting Burden of Healthcare Data Using Robust Principal Component Analysis	Les Servi, Randy Paffenroth, Melanie Jutras, and Deon Burchett
12:30-12:50	Mining Concepts for a COVID Interface Terminology for Annotation of EHRs	Vipina Keloth, Shuxin Zhou, Luke Lindemann, Gai Elhanan, Andrew Einstein, James Geller, and Yehoshua Perl
12:50-1:10	Lung Pattern Classification Via DCNN	Jing Selena He, Meng Han, Lei Yu, and Chao Mei
1:10-1:30	Attention-Based LSTM Network for COVID-19 Clinical Trial Parsing	Xiong Liu, Luca A. Finelli, Greg L. Hersch, and Iya Khalil
1:30-1:50	Challenges and Barriers in Applying Natural Language Processing to Medical Examiner Notes from Fatal Opioid Poisoning Cases	Daniel Harris, Christian Eisinger, Yanning Wang, and Chris Delcher
1:50-2:10	An Empirical Study on Efficiency of a Dictionary Based Viterbi Algorithm for Word Segmentation	Sudhir Aggarwal, Shiva Houshmand, Tathagata Mukherjee, and James Parsons
2:10-2:30	Coffee Break	
2:30-2:50	Clinical Sublanguage Trend and Usage Analysis from a Large Clinical Corpus	Yijun Shao, Guy Divita, Terri Elizabeth Workman, Doug Redd, Jennifer Garvin, and Qing Zeng-Treitler
2:50-3:10	Visualizing Effects of COVID-19 Social Isolation with Residential Activity Big Data Sensor Data	Anuradha Rajkumar, Bruce Wallace, Laura Ault, Julien Larivi�re-Chartier, Frank Knoefel, Rafik Goubbran, Jeff Kaye, and Neil Thomas
3:10-3:30	A Cancelable Multi-Modal Biometric Based Encryption Scheme for Medical Images	Alycia Carey and Justin Zhan
3:30-3:50	Predicting Clinical Deterioration in Hospitals	Laleh Jalali, Hsiu-Khuern Tang, Richard Goldstein, and joaquin Alvarez Rodriguez
3:50-4:10	Examining Deep Learning Models with Multiple Data Sources for COVID-19 Forecasting	Lijing Wang, Aniruddha Adiga, Srinivasan Venkatramanan, Jiangzhuo Chen, Bryan Lewis, and Madhav Marathe
4:10-4:30	A Machine Learning Based Modeling of the Cytokine Storm as it Relates to COVID-19 Using a Virtual Clinical Semantic Network (vCSN)	Abrar Rahman, John Kriak, Rick Meyer, Sidney Goldblatt, and Fuad Rahman
4:30-4:50	Patterns of retinal nerve fiber layer loss in patients with glaucoma identified by deep archetypal analysis	Sidharth Mahotra and Siamak yousefi
4:50-5:10	Analyzing Adverse Event Signal Detection with Publicly Available Web Sources	Alex Salamun, Stephanie Duque, and Praveen Madiraju
5:10	Closing Remarks	

2nd IEEE Workshop on Machine Learning for Big Data Analytics in Remote Sensing

Chairs: Maryam Rahnemoonfar, Grant Scott

Time	Accepted Papers	Presenter
9:00-9:30	Evaluation of Fuzzy Integral Data Fusion Methods for Rare Object Detection in High-Resolution Satellite Imagery	Cannaday
9:30-10:00	Image translation between SAR and optical imagery for Arctic wildfire analysis	Hu
10:00-10:30	A Multi-task Two-stream Spatiotemporal Convolutional Neural Network for Convective Storm Nowcasting	Zhang
10:30-11:00	Enabling Machine-Assisted Visual Analytics for High-Resolution Remote Sensing Imagery with Enhanced Benchmark Meta-Dataset Training of NAS Neural Networks	Hurt
11:00-11:30	Deep Ice Layer Tracking and Thickness Estimation using Fully Convolutional Networks	Varshney
11:30-12:00	Investigating the Relationship Between High-Latitude Dust and Precipitation	Maharaj
12:00-12:30	Comprehensive Semantic Segmentation on High Resolution Aerial Imagery for Natural Disaster Assessment	Chowdhury
12:30-13:00	Extending Deep Convolutional Neural Networks from 3-Color to Full Multispectral Remote Sensing Imagery	Bajkowski

PSBD 2020 Detailed Program Schedule		
Sunday 13 December, 2020		
Time	Title	Presenter/Author
08:55am – 09:00am	Session PSBD20_1: Session Opening Chair: Alfredo Cuzzocrea	
09:00am – 10:15am	Session PSBD20_2: Privacy-Preserving Big Data Methods and Techniques Chair: TBA	
09:00am – 09:15am	Preserving Privacy of Temporal Big Data	Anifat M. Olawoyin, Carson K. Leung, and Alfredo Cuzzocrea
09:15am – 09:30am	Practical Privacy-Preserving Data Science With Homomorphic Encryption: An Overview	Michela Iezzi
09:30am – 09:45am	Privacy-Preserving Multiple Imputation	Wei Qiu, Yangsibo Huang, and Weichen Wu
09:45am – 10:00am	Differentially Private Frequency Sketches for Intermittent Queries on Large Data Streams	Sinan Yildirim, Kamer Kaya, Soner Aydin, and Hakan Bugra Erentug
10:00am – 10:15am	Privacy Preserving Proxy for Machine Learning as a Service	Keshav Kasichainula, Hadi Mansourifar, and Weidong Shi
10:15am – 11:30am	Session PSBD20 3: Advanced Knowledge-Based and Semantics-Based Approaches for Supporting Privacy and Security of Big Data Chair: TBA	
10:15am – 10:30am	Knowledge Graph Anonymization using Semantic Anatomization	Maxime Thouvenot, Olivier Cure, and Philippe Calvez
10:30am – 10:45am	Actionable Knowledge Extraction Framework for COVID-19	Mohammad Masum, Hossain Shahriar, Hisham Haddad, Sheikh Ahamed, Sweta Sneha, Mohammad Rahman, and Alfredo Cuzzocrea
10:45am – 11:00am	Measuring Semantic Similarity across EU GDPR Regulation and Cloud Privacy Policies	Lavanya Elluri, Karuna Pande Joshi, and Anantaa Kotal
11:00am – 11:15am	Cloud-based Encrypted EHR System with Semantically Rich Access Control and Searchable Encryption	Redwan Walid, Karuna Pande Joshi, Seung Geol Choi, and Dae-young Kim
11:15am – 11:30am	Detection of Energy Theft in Smart Grids using Electricity Consumption Patterns	Dabeeruddin Syed, Haitham Abu-Rub, Shady S. Refaat, and Le Xie
11:30am – 12:45pm	Session PSBD20_4: Intelligent Secure Methods and Techniques over Big Data Chair: TBA	
11:30am – 11:45am	z-anonymity: Zero-Delay Anonymization for Data Streams	Nikhil Jha, Thomas Favale, Luca Vassio, Martino Trevisan, and Marco Mellia
11:45am – 12:00pm	Secure and Efficient Trajectory-Based Contact Tracing using Trusted Hardware	Fumiyuki Kato, Yang Cao, and Masatoshi Yoshikawa
12:00pm – 12:15pm	Closed Itemset based Sensitive Pattern Hiding for Improved Data Utility and Scalability	Himanshu Makkar, Durga Toshniwal, and Shalini Jangra
12:15pm – 12:30pm	Exploring the Impact of Resampling Methods for Malware Detection	Paula Branco
12:30pm – 12:45pm	Highly Accurate CNN Inference Using Approximate Activation Functions over Homomorphic Encryption	Takumi Ishiyama, Takuya Suzuki, and Hayato Yamana
12:45pm – 12:50pm	Session PSBD20_5: Session Closing Chair: Alfredo Cuzzocrea	

Second Workshop on Big Data Predictive Maintenance Using Artificial Intelligence <i>Workshop Chairs: Rituparna Datta, Ryan Benton, Aviv Segev</i> Date: 10 December 2020		
Time	Title	Presenter/Author
9:00-9:05	Opening Remarks	
9:05-9:40	Invited Talk	Raju Gottumukkala
9:40-10:05	S37201: Multipath Temporal Convolutional Network for Remaining Useful Life Estimation	Ivan Melendez-Vazquez, Rolando Doelling, and Oliver Bringmann
10:05-10:30	S37202: Autoencoder-based Condition Monitoring and Anomaly Detection Method for Rotating Machines	Sabtain Ahmad, Kevin Styp-Rekowski, Sasho Nedelkoski, and Odej Kao
10:30-10:55	S37203: Predictive Maintenance on Aircraft and Applications with Digital Twin	Sam Heim, Jason Clemens, James Steck, Christopher Basic, David Timmons, and Kourtney Zwiener
10:55-11:20	S37204: Animal Behavior Prediction with Long Short-Term Memory	Henry Roberts and Aviv Segev
11:20-11:40	Coffee Break	
11:40-12:05	N238: Big Data Processing for Power Grid Event Detection	Bruno Paes Leao, Dmitriy Fradkin, Yubo Wang, and Sindhu Suresh
12:05-12:30	BigD632: Dynamic Thresholding Leading to Optimal Inventory Maintenance	Suresh Choubey, Ryan Benton, and Tom Johnsten
12:30-12:55	BigD383: Method for Selecting a Data Imputation Model Based on Programming by Example for Data Analysts	Hiroko Nagashima and Yuka Kato
12:55-1:20	BigD249: A Hybrid Decision Tree-Neural Network (DT-NN) Model for Large-Scale Classification Problems	Jarrod Carson, Kane Hollingsworth, Rituparna Datta, George Clark, and Aviv Segev
1:20-1:30	Closing Remarks	

The 6th International Workshop on Solar & Stellar Astronomy Big Data (SABiD) 2020 <i>Workshop Chairs: Rafal Angryk, Berkay Aydin, Dustin J. Kempton</i>		
Time	Title	Presenter/Author
9:00-9:25 am	S38202: Exploring Predictive Capabilities of GOES and SDO/EVE Data for Flare Forecasting	Jonathan Ehrenguber
9:25-9:50 am	S38205: An Application of the Deep Belief Networks to the prediction of massive Solar Flare Occurrence Using SDO/HMI Data	Ali Abed
9:50-10:15 am	S38204: On The Effectiveness of Imaging of Time Series for Flare Forecasting Problem	Yang Chen
10:15-10:40 am	S38207: First Steps Toward Synthetic Sample Generation for Machine Learning Based Flare Forecasting	Maxwell Hostetter
10:40-10:55 am	15 min. Break	
10:55-11:20 am	S38203: All-Clear Flare Prediction using Interval-based Time Series Classifiers	Anli Ji
11:20-11:45 am	S38206: Deep Neural Network-based Active Region Magnetogram Patch Super Resolution	M. Shuebuddin Habeeb
11:45-12:10 pm	S38201: A Framework for Detecting Polarity Inversion Lines from Longitudinal Magnetograms	Xumin Cai
12:10-12:25 pm	15 min. Break	
12:25-1:10 pm	Invited Talk	Mark Cheung, LMSAL
	Closing Remarks	

Big Data Analytic Technology for Bioinformatics and Health Informatics (KDDDBHI) <i>Workshop Chairs: Xin Deng, Donghui Wu</i>		
Time	Title	Presenter/Author
9:20 – 9:30	Opening Remarks	Dr. Xin Deng
9:30 – 10:30	Keynote Speech: Graph-based Representation Learning for Electronic Health Records	Prof. Edward Choi
10:30 – 10:50	S39205: Assessing SARS-CoV-2 Spike Protein Flexibility for Potential Therapeutic Targets with a Combined Simulation and Deep Learning Approach	Serena Chen
10:50 – 11:10	S39209: Multi-label Detection and Classification of Red Blood Cells in Microscopic Images	Wei Qiu
11:10 – 11:30	S39207: Multimodal Data Representation with Deep Learning for Extracting Cancer Characteristics from Clinical Text	Mohammed Alawad
11:30 – 11:50	S39203: Perception detection using Twitter	Vanja Ljevar
11:50 – 12:10	Coffee Break	
12:10 – 12:30	S39208: A Prototype Application to Identify LGBT Patients in Clinical Notes	Terri Elizabeth Workman
12:30 – 12:50	S39202: Contrast-resolution Evaluation of Fourier Based High Frame Rate Imaging	Zhaohui Wang
12:50 – 13:10	BigD658: Does Yoga Make You Happy? Analyzing Twitter User Happiness using Textual and Temporal Information	Tunazzina Islam
13:10 – 13:30	BigD600: A Deep Recurrent Neural Network to Support Guidelines and Decision Making of Social Distancing	Mohammed Aledhari

Advances in High Dimensional Big Data (AdHD Big Data) <i>(AdHD)Sotiris Tasoulis, Aristidis Vrahatis</i>			
Time	Paper ID	Title	Presenter/Author
9:30 – 9:40		Intro to DPM	
9:40 – 10:05	S40201	Discriminative Pattern Mining for Natural Language Metaphor Generation	Jennifer Brooks
10:05 – 10:20	BigD591	T Discriminative Pattern Mining Workshop (DPM)	Wyatt Green
10:20-10:25		Intro to AdHD	
10:25-10:55	Keynote speech	Scalable distributed multi-dimensional data structures for big data management	Spyros Sioutas
10:55-11:10		Break	
11:10-11:35	S30203	Xinyu Chen, Ian Beaver, and Cynthia Freeman, "Fine-Tuning Language Models For Semi-Supervised Text Mining"	Xinyu Chen
11:35-11:50	S30201	Aikaterini Karampasi, Ioannis Kakkos, Stavros-Theofanis Miloulis, Ioannis Zorzos, Georgios Dimitrakopoulos, Kostakis Gkiatis, Panteleimon Asvestas, and George Matsopoulos, "A Machine Learning fMRI Approach in the Diagnosis of Autism"	Aikaterini Karampasi
11:50-12:15	S30202	Ivan Kyosev, Iulia Paun, Yashar Moshfeghi, and Nikos Ntarmos, "Measuring Distances Among Graphs En Route To Graph Clustering"	Ivan Kyosev
12:15-12:40	S30206	Panagiotis Anagnostou, Petros Barbas, Aristidis G. Vrahatis, and Sotiris K. Tasoulis, "Approximate kNN Classification for Biomedical Data"	Petros Barbas
12:40-1:05	S30205	Georgios Drakopoulos, Ioanna Giannoukou, Phivos Mylonas, and Spyros Sioutas, "A Graph Neural Network For Assessing The Affective Coherence Of Twitter Graphs"	Georgios Drakopoulos
1:05-1:15		Closing	

Discriminative Pattern Mining Workshop (DPM 2020) And

4th International Workshop on Big Data Analytics for Cyber Intelligence and Defense (BDA4CID 2020)		
Workshop Chairs: Ryan Benton, Tom Johnsten, Suresh Choubey, Stephen McGough and Amir Atapour Abarghouei		
Time	Title	Presenter/Author
	Session 1: Discriminative Pattern Mining Workshop	
9:00-9:05am	Intro to Discriminative Pattern Mining Workshop	Ryan Benton
9:05-9:30am	S40201: Discriminative Pattern Mining for Natural Language Metaphor Generation	Jennifer Brooks and Abdou Youssef
9:30-10:00am	BigD591: TADS: Transformation of Anomalies in Data Streams	Wyatt Green, Tom Johnsten, and Ryan Benton
10:00-10:15am	Discriminative Pattern Mining Workshop Closing Remarks	Ryan Benton
10:15-10:45am	Break	
	Session 2: 4th International Workshop on Big Data Analytics for Cyber Intelligence and Defense	
10:45-10:50am	Intro to Workshop on Big Data Analytics for Cyber Intelligence and Defense	Stephen McGough
10:50-11:20am	S34201: FlightSense: A Spoofer Detection and Aircraft Identification System using Raw ADS-B Data	Nikita Joseph, Chaity Banerjee, Eduardo Pasiliao, and Tathagata Mukherjee
11:20-11:50am	S34202: Rumor Detection on Social Networks: A Sociological Approach	Neelam Jogalekar, Vahida Attar, and Girish Palshikar
11:50-12:20pm	S34203: Resolving the cybersecurity Data Sharing Paradox to scale up cybersecurity via a co-production approach towards data sharing	Amir Atapour-Abarghouei, Stephen McGough, and David Wall
12:20-12:35pm	BDA4CID Workshop Closing Remarks	Stephen McGough

Workshop on Deviant Activities on Social Media		
<i>Workshop Chairs: TBD</i>		
Time	Title	Presenter/Author
1:30pm - 1:40pm	Welcome	
1:40pm - 2:40pm	Keynote	Dr. Amy Bruckman
2:40pm - 3:40pm	Track 1 - Deviant Activities Detection	
2:40pm - 3:00pm	S41205 - Toward A Multilingual and Multimodal Data Repository for COVID-19 Disinformation	Yichuan Li, Bohan Jiang, Kai Shu, and Huan Liu
3:00pm - 3:20 pm	BigD580 - Illicit Activity Detection in Large-Scale Dark and Opaque Web Social Networks	Dhara Shah, Timothy Harrison, Christopher Freas, David Maimon, and Robert Harrison
3:20pm - 3:40pm	BigD663 - How biased are American media outlets? A framework for presentation bias regression	Minh Tran
3:40pm - 3:50pm	Coffee Break	
3:50pm - 4:50pm	Track 2 - Deviant Information Operations and Campaigns	
3:50pm - 4:10pm	S41201 - Understanding Online Information Operations: Development of an Influence Network for Scientific Inquiry Testing Environment (INSITE)	Courtney Crooks, Matthew Canham, Tom McNeil, David Muchlinski, and Ben Sawyer
4:10pm - 4:30pm	BigD554 - Unveiling Ideological Trends Through Data Analytics to Construe National Security Instabilities	PEDRO CARDENAS CANTO, Georgios Theodoropoulos, Boguslaw Obara, and IBAD KURESHI
4:30pm - 4:50pm	S41203 - Profiling Astroturfing Facebook Users During Three Contiguous Israeli Election Periods	Jonathan Schler, Elisheva Bonchek Dokow, Tomer Vainstein, Moshe Gotam, and Mike Teplitsky
4:50pm - 5:50pm	Track 3 - Deviance on Twitter	
4:50pm - 5:10pm	S41202 - Pro-ISIS Tweets Analysis Using Machine Learning Techniques	Julia Thee, Izzat Alsmadi, and Samer Al-khateeb
5:10pm - 5:30pm	S41204 - How to Control Coronavirus Conspiracy Theories in Twitter? A Systems Thinking and Social Networks Modeling Approach	Mustafa Alassad, Muhammad Nihal Hussain, and Nitin Agarwal
5:30pm - 5:50pm	BigD382 - Collecting Domain-Relevant Tweets: Creation and Application of a New Framework	Alina Campan, Traian Marius Truta, Shawn Huesman, Celine Wardrop, and Alyssa Appelman
5:50pm - 6pm	Closing Remarks	

2020 Innovation Workshop on Transforming Big Data into Actionable Knowledge (BiDAW 2020) December 12th, 2020 <i>Workshop Chairs: Mahdi Bohlouli, Maria-Esther Vidal, Ali Behravan</i>		
Time	Title	Presenter/Author
09:00 – 09:25	S45201: Building A Non-Personalized Recommender System by Learning Product and Basket Representation	savaş yıldırım, Şebnem Güneş Söyler, and Özgür Akarsu
09:25 – 09:50	S45: ReBERT: Retraining BERT on retail domain for task-oriented dialogue systems	Pratik Jayarao and Arpit Sharma
09:50 - 10:15	S45204: Scalable Multi-Criteria Decision-Making: A MapReduce Deployed Big Data Approach	Mahdi Bohlouli and Martin Schrage
10:15 - 10:40	S45205: Cloud enabled Multi-modal Knowledge Modeling for Learning and Skill Management (KaaS)	Mahdi Bohlouli and Sebastian Hellekes
10:40 - 11:20	Keynote Speech, S45203: Preparing the ground for AI with heterogeneous data integration leveraging Knowledge Graphs	Prof. Dr. Sören Auer
11:20 – 11:30	Closing Remarks	

User Understanding from BigData <i>Workshop Chairs: Claire Ding, Wutao Wei, Ying Lu</i>		
Time (Est)	Title	Presenter/Author
11:30 – 11:40 am	Opening Remarks	Claire Ding, Wutao Wei, Ying Lu
11:40 – 12:10 pm	S22202: Representation of Click-Stream Data Sequences for Learning User Navigational Behavior by Using Embeddings	Erdi Olmezogullari and Mehmet Aktas
12:10-12:20 pm	Coffee Break	
12:20-12:50 pm	S22206: Understanding User Understanding: What do Developers Expect from a Cognitive Assistant?	Glauca Melo, Edith Law, Paulo Alencar, & Donald Cowan
12:50 – 1:20 pm	S22207: On Variational Inference for User Modeling in Attribute-Driven Collaborative Filtering	Venugopal Mani, Ramasubramanian Balasubramanian, Sushant Kumar, Abhinav Mathur & Kannan Achan
1:20 – 1:30 Closing Remarks		

Seventh International Workshop on High Performance Big Graph Data Management, Analysis, and Mining (BigGraphs 2020)

Workshop Chairs: Nesreen Ahmed, Mohammad Al Hasan, Shaikh Arifuzzaman, Kamesh Madduri

Time	Title	Presenter/Author
9:00-9:05 am	Opening Remarks	
9:05-9:30 am	S19203 Approximately and Efficiently Estimating Dynamic Point-to-Point Shortest Path	Alok Tripathy and Oded Green
9:30-9:55 am	BigD319 Agent-Navigable Dynamic Graph Construction and Visualization over Distributed Memory	Justin Gilroy, Satine Paronyan, Jonathan Acoltzi, and Munehiro Fukuda
9:55-10:20 am	BigD448 Graph Filtering to Remove the "Middle Ground" for Anomaly Detection	William Eberle and Lawrence Holder
10:25-11:15 am	Keynote	Madhav Marathe
11:20-11:45 am	S19205 Massively Parallel Processing Database for Sequence and Graph Data Structures Applied to Rapid-Response Drug Repurposing	Chris Rickett, Kristyn Maschoff, and Sreenivas Sukumar
11:45 am-12:10 pm	BigD559 Retrieving Entities from Knowledge Graphs without Knowing Much: On Learning Generalizable Patterns and Indexing	Yuting Xie and Tingjian Ge
12:10-12:35 pm	S19206 Graph Force Learning	Ke Sun, Jiaying Liu, Shuo Yu, Bo Xu, and Feng Xia
12:35-1:00 pm	S19207 Memory Efficient Graph Convolutional Network based Distributed Link Prediction	Damitha Senevirathne, Isuru Wijesiri, Suchitha Dehigaspitiya, Miyuru Dayarathna, Sanath Jayasena, and Toyotaro Suzumura
1:10-1:35 pm	BigD579 Understanding Coarsening for Embedding Large-Scale Graphs	Taha Atahan Akyıldız, Amro Alabsi Alundi, and Kamer Kaya
1:35-2:00 pm	BigD336 Learning Embeddings of Directed Networks with Text-Associated Nodes---with Application in Software Package Dependency Networks	Kexuan Sun, Shudan Zhong, and Hong Xu
2:00-2:25 pm	BigD434 Relational Graph Embeddings for Table Retrieval	Mohamed Trabelsi, Zhiyu Chen, Brian D. Davison, and Jeff Heflin
2:25-2:30 pm	Closing Remarks	

First International Workshop on the Internet of Things Data Analytics (IoTDA)

Workshop Chairs: Eyhab Al-Masri (University of Washington Tacoma), Chi-Hua Chen (Fuzhou University, China)

Time	Title	Presenter/Author
10:00 am – 10:10 am	Opening Remarks and Welcome	
10:15 am – 10:30 am	(S28201) Pedestrian Trajectory Prediction Using Pre-trained Machine Learning Model for Human-Following Mobile Robot	Rina Akabane and Yuka Kato
10:35 am – 10:50 am	(S28204) Elucidating the Extent by Which Population Staying Patterns Help Improve Electricity Load Demand Predictions	Guillaume Habault, Shinya Wada, Rui Kimura, and Chihiro Ono
10:55 am – 11:10 am	(S28212) Failure Prediction in Datacenters Using Unsupervised Multimodal Anomaly Detection	Minglu Zhao, Reo Furuhashi, Mulya Agung, Hiroyuki Takizawa, and Tomoya Soma
11:15 am – 11:30 am	(S28217) Understanding Bit-Error Trade-off of Transform-based Lossy Compression on Electrocardiogram Signals	Aekyeung Moon, Seung Woo Son, Jiuk Jung, and Yun Jeong Song
11:35 am – 11:50 am	(BigD748) Automatic Device Identification and Anomaly Detection with Machine Learning Techniques in Smart Factories	Chin-Wei Tien, Tse-Yung Huang, Ping Chun Chen, and Jenq-Haur Wang
12:00 pm – 1:00 pm	Break	
1:00 pm – 1:15 pm	(S28205) Towards AIOps in Edge Computing Environments	Soeren Becker, Florian Schmidt, Anton Gulenko, Alexander Acker, and Odej Kao
1:20 pm – 1:35 pm	(S28208) Synchronized Preprocessing of Sensor Data	Amal Tawakuli, Daniel Kaiser, and Thomas Engel
1:40 pm – 1:55 pm	(S28214) A Scalable and Dependable Data Analytics Platform for Water Infrastructure Monitoring	Felix Lorenz, Morgan Geldenhuys, Harald Sommer, Frauke Jakobs, Carsten LÄ¼ring, Volker Skwarek, Ilja Behnke, and Lauritz Thamsen
2:00 pm – 2:15 pm	(S28206) Anomaly and Degradation Detection Using Subspace Tracking in Streaming Data	Kyungduck Cha, Carol Sadek, and Zohreh Asgharzadeh
2:20 pm – 2:35 pm	(S28215) Real-Time Machine Learning for Air Quality and Environmental Noise Detection	Sayed Khushal Shah, Zeenat Tariq, Jeehwan Lee, and Yugyung Lee
2:40 pm – 2:55 pm	(S28209) A Reference Model for IoT Embodied Agents Controlled by Neural Networks	Nathalia Nascimento, Paulo Alencar, Donald Cowan, and Carlos Lucena
2:55 pm – 3:15 pm	Break	
3:20 pm – 3:35 pm	(S28218) Automatic Multimodal Heart Disease Classification using Phonocardiogram Signal	Zeenat Tariq, Sayed Khushal Shah, and Yugyung Lee
3:40 pm – 3:55 pm	(S28220) Large-scale Data Integration for Facilities Analytics: Challenges and Opportunities	Balaje T. Thumati, Halasya Siva Subramania, Rajeev Shastri, Karthik Kalyana Kumar, Nicole Hessner, Vincent Villa, Aaron Page, and David Followell
4:00 pm – 4:15 pm	(S28202) Towards Policy-aware Edge Computing Architectures	Pratik Baniya, Gaurav Bajaj, Jerry Lee, Clifton Francis, Ardesheer Bastani, and Mahima Agumbe Suresh
4:20 pm – 4:35 pm	(S28216) Enhancing Resource Provisioning Across Edge-based Environments	Eyhab Al-Masri and James Olmsted
4:40 pm – 4:45 pm	Closing Remarks	

International Workshop on Data Analytics for Smart Health

Workshop Chairs: Harishchandra Dubey, Xiaoqian Jiang, Arindam Pal

Time	Title	Presenter/Author
10:45- 11 AM	Welcome from Chairs	
11 AM - 12 PM	Keynote1: Smart Sensing and Analytics for Cognitive Health Care	Prof. Sajal K. Das
12-12:30 PM	Paper ID S27206: Utilizing Social Media for Identifying Drug Addiction and Recovery Intervention	Saptarshi Ghosh
12:30 -1 PM	Paper ID BigD704: Privacy preserving time-series forecasting of user health data streams	Vladimir Vlassov
1-1:30 PM	Paper ID S27208: Google Trends Analysis of COVID-19 Pandemic	Zhenhe Pan
1:30 - 2:30 PM	Keynote2: Digital Behavior Markers: The Beginning of a Digital Twin	Prof. Diane J. Cook
2:30 - 3:30 PM	Keynote3	Prof. Lucila Ohno-Machado
3:30 - 4:30 PM	Keynote4: Machine Reading for Precision Medicine	Dr. Hoifung Poon
4:30 - 5:00 PM	Paper ID S27207: A Framework for Edge-Assisted Healthcare Data Analytics using Federated Learning	Suprio Ray
5:00 - 5:30 PM	Paper ID BigD525: Not All Areas Are Equal: Detecting Thoracic Disease with ChestWNet	Fang Jin
5:30 – 5:40 PM	Closing Remarks	

Special Sessions

6th SPECIAL SESSION ON INTELLIGENT DATA MINING

After the successes of the first, second, third, fourth and fifth editions of Special Session on Intelligent Data Mining in Santa Clara, CA (2015), Washington, DC (2016), Boston, MA (2017), Seattle, WA, (2018), Los Angeles, CA, (2019) and the sixth Special Session on Intelligent Data Mining in Atlanta, GA (2020) will continue promoting and disseminating the knowledge concerning several topics and technologies related to data mining science.

Artificial Intelligence (AI) & Machine Learning (ML) fields are interdisciplinary, including computer science, mathematics, psychology, linguistics, philosophy, neuroscience etc. This interdisciplinary special session seeks scientific understanding on data and intelligence.

This session may help to create scientific evolution to propose robust and powerful schemes between human nature and big data processing.

Session Date: December 11, 2020

6 th Special Session on Intelligent Data Mining <i>Workshop Chair: Uraz Yavanoglu</i>		
Time	Title	Presenter/Author
09:00-09:15	BigD681-Exploiting Hybrid Subword Information for Chinese Historical Named Entity Recognition	Chengxi Yan and Jun Wang
09:15-09:30	BigD734-Urban Heat Islands: Beating The Heat With Multi-Modal Spatial Analysis	Marcus Yong and Kwan Hui Lim
09:30-09:45	BigD430-Distributed Mining of Spatial High Utility Itemsets in Very Large Spatiotemporal Databases using Spark In-Memory Computing Architecture	Uday Kiran Rage , Sadanori Ito, Minh-Son Dao, Koji Zettsu, Cheng-wei Wu, Yutaka Watanobe, Incheon Paik, and Truong Cong Thang
09:45-10:00	BigD372-EvaNet: An Extreme Value Attention Network for Long-Term Air Quality Prediction	Zechuan Chen , Haomin Yu, Yangli-ao Geng, Qingyong Li, and Yingjun Zhang
10:00-10:15	BigD363-ModelRecycling: Predicting user's interest with connected predictive models	Yoshiki Matsune , Kota Tsubouchi, and Nobuhiko Nishio
10:15-10:30	BigD342-Discovering Closed Periodic-Frequent Patterns in Very Large Temporal Databases	Likhitha Palla, Ravikumar P, Uday Kiran Rage , Yuto Hayamizu, Kazuo Goda, Masashi Toyoda, Koji Zettsu, and Sourabh Shrivastava
10:30-10:45	BigD205-An Approximation Method for Large Graph Similarity	Danfeng Zhao, Zhou Huang , Feng Zhou, Antonio Liotta, and Dongmei Huang
10:45-11:00	SP02209-Multi-source Machine Learning for AQI Estimation	Quoc Dat Duong, Quang Le, Tan Loc Nguyen Tai, Quang Dong Bo, Dat Nguyen, Minh-Son Dao , and Binh Nguyen
11:00-11:15	Coffee Break	
11:15-11:30	SP02213-Explainable Software vulnerability detection based on Attention-based Bidirectional Recurrent Neural Networks	Yi Mao
11:30-11:45	BigD650-Towards Industrial Internet of Things in Steel Manufacturing: A Multiple-Factor-based Detection System of Longitudinal Surface Cracks	Fucun Li , Ailing Yang, Huaming Chen, Geng Sun, Fang Wang, Yunzhou Xie, Jiaqi Li, and Jun Shen
11:45-12:00	BigD730-A Comparative Evaluation of Top-N Recommendation Algorithms: Case Study with Total Customers	Idir Benouaret and Sihem Amer-Yahia
12:00-12:15	BigD703-On Convolutional Autoencoders to Speed Up Similarity-Based Time Series Mining	Yuri Gabriel Aragão da Silva and Diego Furtado Silva
12:15-12:30	BigD405-On the Evaluation and Deployment of Machine Learning Approaches for Intrusion Detection	Felix Heine, Tim Laue , and Carsten Kleiner
12:30-12:45	SP02207-The Best of Both Worlds: Context-Powered Word Embedding Combinations for Longitudinal Text Analysis	Laura Rettig , Regula Hänggli, and Philippe Cudré-Mauroux
12:45-13:00	SP02212-Finding Overlapping Clusters in a Highly Connected Graph from a Given Difference Density	Kittichai Lavangnananda and Muhammad Aslam

13:00-13:15	SP02215-Intelligent Chatter Detection in Milling using Vibration Data Features and Deep Multi-Layer Perceptron	Batihan Sener , Gokberk Serin, Ugur Gudelek, Murat Ozbayoglu, and Ozgur Unver
1:15-1:30	Coffee Break	
13:30-13:45	SP02219-Speculator and Influencer Evaluation in Stock Markets by Using Social Media	Mustafa Dogan, Omer Metin, Elif Tek, Semih Yumusak, and Kasim Oztoprak
13:45-14:00	SP02220-Extraction of Product Defects and Opinions from Customer Reviews by Using Text Clustering and Sentiment Analysis	Mustafa Cataltas, Sevcen Dogramaci, Semih Yumusak, and Kasim Oztoprak
14:00-14:15	SP02221-Named Entity Recognition on Morphologically Rich Language: Exploring the Performance of BERT with varying Training Levels	Yuksel Pelin Kilic, Duygu Dinc, and Pinar Karagoz
14:15-14:30	SP02222-Sentiment Analysis for Turkish Unstructured Data by Machine Translation	Mustafa Yıldırım, Feyza Yıldırım Okay, and Suat Özdemir
14:30-14:45	SP02223-A Novel Weighted FP-Stream Algorithm for IoT Data Streams	Halil Ibrahim Dede, Cemile Timurkaan, Metehan Guzel, and Suat Ozdemir
14:45-15:00	BigD691-An experimental study of different machine and deep learning techniques for classification of encrypted network traffic	ThankGod Obasi and M. Omair Shafiq
15:00-15:15	BigD387-Accelerated Evaluation of Autonomous Drivers using Neural Network Quantile Generators	Edward Schwalb
15:15-15:30	Coffee Break	
15:30-15:45	SP02201-Stock Price Prediction Under Anomalous Circumstances	Wei Wu , Jinlong Ruan, and Jiebo Luo
15:45-16:00	SP02202-Content-based Analysis of the Cultural Differences between TikTok and Douyin	Li Sun , Haoqi Zhang, Songyang Zhang, and Jiebo Luo
16:00-16:15	SP02214-Face Off: Polarized Public Opinions on Personal Face Mask Usage during the COVID-19 Pandemic	Neil Yeung , Jonathan Lai, and Jiebo Luo
16:15-16:30	SP02216-Item-Based Collaborative Filtering and Association Rules for a Baseline Recommender in E-Commerce	Jessica Lourenco and Aparna Varde
16:30-16:45	SP02204-Recipe for Creating a Highly Accurate Wake Word Engine	Buvaneswari Ramanan , Lawrence Drabeck, Thomas Woo, Troy Cauble, and Anil Rana
16:45-17:00	SP02205-A Sentimental and Semantical Analysis on Facebook Comments to Detect Latent Patterns	Mahdi Naser Moghadasi , Zohreh Safari, and Yu Zhuang
17:00-17:15	SP02208-Local Outlier Detection for Multi-type Spatio-temporal Trajectories	Xumin Cai , Berkay Aydin, Saurabh Maydeo, Anli Ji, and Rafal Angryk
17:15-17:30	SP02210-Uncovering the Folding Landscape of RNA Secondary Structure Using Deep Graph Embeddings	Egbert Castro , Andrew Benz, Alexander Tong, Guy Wolf, and Smita Krishnaswamy
17:30-17:45	SP02211-Solar Line-of-Sight Magnetograms Super-Resolution Using Deep Neural Networks	Mohammed Shuebuddin Habeeb , Berkay Aydin, Azim Ahmadzadeh, Manolis Georgoulis, and Rafal Angryk
17:45-18:00	Coffee Break	
18:00-18:15	SP02217-An evaluation of machine learning methods for domain name classification	Amit Garg , Nachiket Trivedi, Junlan Lu, Magdalini Eirinaki, Bin Yu, and Femi Olumofin
18:15-18:30	SP02218-Higher-order Link Prediction Using Triangle Embeddings	Neeraj Chavan and Katerina Potika
18:30-18:45	BigD227-Method for Customizable Automated Tagging: Addressing the problem of over-tagging and under-tagging text documents.	Maharshi Pandya , Jessica Reyes, and Bob Vanderheyden
18:45-19:00	BigD248-An Automatic Classification of the Primary and the Corresponding Authors in Research Articles	Sukhwan Jung , Rituparna Datta, and Aviv Segev
19:00-19:15	BigD263-IFGAN: Missing Value Imputation using Feature-specific Generative Adversarial Networks	Wei Qiu , Yangsibo Huang, and Quanzheng Li
19:15-19:30	BigD318-Correlation-Based Analytics of Time Series Data	Ramoza Ahsan , Rodica Neamtu, Muzammil Bashir, Elke Rundensteiner, and Gabor Sarkozy
19:30-19:45	BigD432-Sent2Vec: A New Sentence Embedding Representation With Sentimental Semantic	Mahdi Naser Moghadasi and Yu Zhuang
19:45-20:00	BigD706-A Real-Time Whole Page Personalization Framework for E-Commerce	Aditya Mantha , Anirudha Sundaresan, Shashank Kedia, Yokila Arora, Shubham Gupta, Gaoyang Wang, Praveenkumar Kanumala, Stephen Guo, and Kannan Achan
	Closing Remarks	

3rd Special Session on HealthCare Data

Health data differs from other industries' data in terms of structure, context, importance, volatility, availability, traceability, liquidity, change speed, usage and sources from which it is collected. As medicine is a constantly developing science, healthcare sector also. In this new emerging research area which stands at the intersection of several different discipline such as Medicine, Behavioral Science, Supply Chain Management or Big Data Analytics, techniques, methods, applications and devices are continuously developed to be used for the acquisition, storage, processing, analysis, standardization and optimization of every process in the health sector. As the healthcare sector is so challenging and related data are consistently explosive, healthcare organizations are focusing to become smarter in order to overcome the industry's inefficiencies to improve quality of care. "To become smarter" requires impeccable data analytics. All stakeholders in the sector should reveal the deep value of this valuable data in order to apply insights to improve quality of care, clinical outcomes and deliver personalized healthcare value, while reducing medical costs, collaborate across care settings to deliver integrated, personalized care experiences, prevent disease, promote wellness and manage care, build flexibility into operations to support cost reduction and excellence in clinical and business performance and practices.

The general purpose of this special session in IEEE BigData 2020 conference is to bring together researchers, academicians and sector employees from different fields and disciplines and provide them an independent platform to exchange information on their researches, ideas and findings about healthcare data and its analytics. It is also aimed to encourage debate on how big data can effectively support the healthcare in terms of diagnosis, treatment and population health, and to develop a common understanding for research conducted in this multidisciplinary field.

3rd Special Session on HealthCare Data		
Chair: Sultan Turhan, Ozgun Pinarer		
Time	Title	Presenter/Author
09:00-09:10	Welcome	
09:10-09:35	BigD309: Prediction of Nocturia in Live Alone Elderly Using Unobtrusive In-Home Sensors	Barry Nuqoba
09:35-10:00	SP03204: Hydra: Cancer Detection Leveraging Multiple Heads and Heterogeneous Datasets	Giuseppe Cuccu
10:00-10:25	SP03208: Transfer learning for decision support in Covid-19 detection from a few images in big data	Divyadharshini Karthikeyan
10:25-10:50	SP03203: Adaptive Anomaly Detection for Dynamic Clinical Event Sequences	Olufemi A. Omitaomu
10:50-11:10	Coffee Break	
11:10-11:35	SP03212: Pandemic Effect: Degradation of Speech Reception Due to Medical Masks	Sultan Turhan, Ozgun Pinarer
11:35-12:00	BigD207: Integrated Time Series Summarization and Prediction Algorithm and its Application to COVID-19 Data Mining	Mogens Graf Plessen
12:00-12:25	SP03207: Development of an Explainable Prediction Model of Heart Failure Survival by Using Ensemble Trees	Pedro A. Moreno-Sanchez
12:25-12:50	SP03201: Stratification of, albeit Mathematical Optimization and Artificial Intelligent (AI) Driven, High-Risk Elderly Outpatients for priority house call visits - a framework to transform healthcare services from reactive to preventive	Chandrasekar Vuppapapati
12:50-14:00	Lunch Break	
14:00-14:25	SP03206: Deep Learning to Predict Hospitalization at Triage: Integration of Structured Data and Unstructured Text	Mahmoud Elbattah
14:25-14:50	SP03209: Investigating Transfer Learning of Smartphone-Sensed Stress in University Populations	Emmanuel Agu
14:50-15:15	SP03210: DeepSEAS: Smartphone-based Early Ailment Sensing Using Coupled LSTM Autoencoders	Shreesha Narasimha Murthy
15:15-15:40	SP03205: INTOSIS: Interactive Observation of Smartphone Inferred Symptoms for In-The-Wild Data	Hamid Mansoor
15:40-16:00	Coffee Break	
16:00-16:25	BigD682: Factors Influencing Human Mobility During The COVID-19 Pandemic in Selected Countries of Europe and North America	Maryam Hosseini
16:25-16:50	BigD210: Orchestration of Thick Data Analytics Based on Conversational Workflows in Healthcare Community of Practice	Jinan Fiaidhi
16:50-17:15	BigD643: Predicting Opioid Overdose Readmission and Opioid Use Disorder with Machine Learning	Praveen Madiraju
17:15-17:30	Closing Remarks	

Explainable Artificial Intelligence in Safety Critical Systems

Dramatic success in machine learning has led to a new wave of artificial intelligence applications that offer extensive benefits to our daily lives. AI is finding its way into a broad range of industries such as education, manufacturing, law enforcement, healthcare, and finance. We are finding that the need to trust these AI based systems with all manner of decision making and predictions is paramount, especially for safety critical systems. However, this is limited by the machine's current inability to explain their decisions and actions to human users.

In order to explore how the AI research community can handle the lack of explainability and trust for AI systems, we organize this special track on Explainable Artificial Intelligence (XAI) in Safety Critical Systems.

XAI aims to create a suite of machine learning techniques that: produce more explainable models, while maintaining a high level of learning performance, extract comprehensible knowledge and actionable plans from AI systems, and enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners.

The special track will focus on technological framework that allows the adoption of explainable artificial intelligence (XAI) in safety-critical systems, which are required to meet governance criteria concerning safety, transparency, accountability, responsibility, and fairness for multiple application areas such as finance, business, and healthcare.

Explainable Artificial Intelligence in Safety Critical Systems <i>General Chair:</i> <i>Qiang Yang, WeBank and the Hong Kong University of Science and Technology</i> <i>PC Chair:</i> <i>Yixin Chen, Washington University in St Louis</i> <i>Lixin Fan, WeBank</i>		
Time	Title	Presenter/Author
9.00-9.10	Opening Remark	Dr. Lixin Fan
9.10-9.40	TBD	Invited Speaker: Prof. Yongfeng Zhang
9.40-10.00	Interpretable feature subset selection: A Shapley value based approach	SP06201
10.00-10.20	Interpreting and Evaluating Black Box Models in a Customizable Way	SP06211
10.20-10.40	Predicting Liquidity Ratio of Mutual Funds via Ensemble Learning	SP06207
10.40-10.50	Link Prediction Based on Heuristics and Graph Attention	SP06210
10.50-11.00	An Evaluation of Pairs Trading in Commodity Futures Markets	SP06203
11.00-11.20	Coffee Break	
11.20-11.50	TBD	Invited Speaker: TBD
11.50-12.10	Encoding Broad Learning System: An Effective Shallow Model For Anti-fraud	SP06204
12.10-12.30	Explainable Machine Learning for Regime-Based Asset Allocation	SP06205
12.30-12.40	A Novel Lasso Regression Model for Sector Rotation Trading Strategies with "Economy-Policy" Cycles	SP06208
12.40-12.50	Exchange Traded Fund Clustering via Metric Learning	SP06209
12.50-1.00	Utilizing Macroeconomic Factors for Sector Rotation based on Interpretable Machine Learning and Explainable AI	SP06206
1.00-1.10	Closing Remark: Prof. Yixin Chen	

Information Granulation in Data Science and Scalable Computing

Granular Computing is a general computation approach for effectively using information granules such as classes, clusters, sets, groups, intervals and networks to build an efficient computational model for complex applications with Big Data, such as huge amounts of data, information and knowledge. Though the label is relatively recent, the notions and principles of Granular Computing and Information Granulation, under different names, have appeared in many related fields, such as information hiding in programming, granularity in artificial intelligence, divide and conquer paradigms in theoretical computer science, interval computing, cluster analysis, fuzzy and rough set systems, neuro computing, quotient spaces, belief functions, approximate analytics, approximate computing, deep neural networks and many others. Thus, "Information granule = Fundamental Piece of Human Knowledge": As a meta-mathematical methodology, computation of information granules gives a theoretical framework of big data analytics.

Special Session on Information Granulation in Data Science and Scalable Computing will continue to address the theory and practice of computation of information granules. It will provide researchers from universities, laboratories and industry with the means to present state-of-the-art research results and methodologies for information granules. The session will also make it possible for scientists and developers to highlight their new research directions and new interactions with novel computing models for information granulation.

The session will focus particularly on currently important research tracks such as social network computing, cloud computing, cyber-security, data mining, machine learning, knowledge management, AI-based systems, intelligent systems and soft computing (neural networks, fuzzy systems, evolutionary computation, rough sets, self-organizing systems), e-Intelligence (Web intelligence, semantic Web, Web informatics), business informatics, bioinformatics and medical informatics.

Information Granulation in Data Science and Scalable Computing		
<i>Special Session Chairs: Shusaku Tsumoto, Dominik Slezak, Tzung-Pei Hong and Shyue-Liang Wang</i>		
Time	Title	Presenter/Author
0900-0925	SP05202 Action Recognition Using Dynamic Mode Decomposition for Temporal Representation 1	Kyle Pawlowaski, Sumit Chakravarty, Ying Xie, and Arjun Joginipelly
0925-0950	SP05201 Extracting Multi-Scale Rotation-Invariant Features in Convolution Neural Networks	Tzung-Pei Hong, Ming-Jhe Hu, Tank-Kai Yin, and Shyue-Liang Wang
0950-1015	SP05212 Deep Pose Alignment	Linh Le, Ying Xie, Saisangaramaleengam Alagapan, Sumit Chakravarty, Pablo Ordonez, Michael Hales, and John Johnson
1015-1040	Coffee Break	
1040-1105	SP05204 Fuzzy High-Utility Pattern Mining based on the Hadoop Framework	Jimmy Ming-Tai Wu, Gautam Srivastava, Min Wei, and Jerry Chun-Wei Lin
1105-1130	SP05205 High-Utility Pattern Mining in Hadoop Environments	Jimmy Ming-Tai Wu, Min Wei, Gautam Srivastava, and Jerry Chun-Wei Lin
1130-1155	SP05207 OSUMI: On-Shelf Utility Mining from Itemset-based Data	Jiahui Chen, Xu Guo, Wensheng Gan, Chien-Ming Chen, Weiping Ding, and Guoting Chen
1155-1230	SP05210 TopHUI: Top-k high-utility itemset mining with negative utility	Wensheng Gan, Shicheng Wan, Jiahui Chen, Chien-Ming Chen, and Lina Qiu
1230-1400	Lunch Break	
1400-1425	SP05203 Mining High-Utility Sequential Patterns in Uncertain Databases	Jerry Chun-Wei Lin, Gautam Srivastava, Yuanfa Li, Tzung-Pei Hong, and Shyue-Liang Wang
1425-1450	SP05209 Order Trajectory Analysis in Hospital Information System	Shusaku Tsumoto, Tomohiro Kimura, and Shoji Hirano
1450-1515	SP05208 Automated Dual Clustering for Clinical Pathway Mining	Shusaku Tsumoto, Tomohiro Kimura, and Shoji Hirano
1515-1540	SP05211 Reinventing Infobright's Concept of Rough Calculations on Granulated Tables for the Purpose of Accelerating Modern Data Processing Frameworks	Mateusz Wnuk, Sebastian Stawicki, and Dominik Slezak
1540-1550	Coffee Break	
1550-1615	SC01204 Data Science in Support Ticket Escalation: IEEE BigData 2020 Cup Challenge	Guohua Hao and Robert Wojciechowski
1615-1640	SC01203 Predicting Escalations in Customer Support: Analysis of Data Mining Challenge Results	Dominik Slezak and Andrzej Janusz
1640-1705	SC01201 Predicting Escalations in Customer Support with Gradient Boosting at the IEEE BigData 2020 Cup	Peter Klimov and Vladimir Funtikov
1705-1730	SC01202 A Hybrid Machine Learning Framework of Gradient Boosting Decision Tree and Sequence Model for Predicting Escalation in Customer Support	Shubham Gupta
1730-1740	Closing Remarks	

The Special Session on Data Marketing as Cross-disciplinary Data Exchange and Collaboration (CDEC2020)

The recent social movement of big data and artificial intelligence has resulted in a tremendous increase in the importance of data. In view of these expectations, there are the externalizations of interdisciplinary issues. Many papers about data utilization have been published, and several approaches for analyzing data have been shared widely. However, there are only a limited number of studies on the process of cross-disciplinary data exchange and collaboration and its ecosystem. Since this process encompasses various activities of different stakeholders, it is difficult to evaluate the patterns or processes quantitatively. Moreover, in the data market, there are not only big data but also small data necessary for decision making. It is essential to discuss the dynamics of such a network of heterogeneous data in different fields.

To address these gaps, we propose to hold a special session to discuss the processes and interactions among data, humans, and society –Data Marketing as Cross-disciplinary Data Exchange and Collaboration (CDEC). The topics taken up at CDEC will involve practical issues such as the analytical tasks performed using data, solutions for challenging social issues, and cross-disciplinary data collaboration and its process. Our special session will target not only cleanly formatted homogenous data, but also heterogeneous data that affect human behaviors, thoughts, and intentions across different domains. We will also focus on a discussion to obtain tacit knowledge of artificial intelligence and data mining by analysis and synthesis. In addition to these research fields, we will attempt to take a cognitive approach toward observing the processes of knowledge discovery and data exchange. It is expected that conflicts and inconsistencies may arise owing to differences in opinion when stakeholders from different knowledge domains have discussions on data-driven decision making. We believe that this special session focusing on the process of cross-disciplinary data exchange and collaboration will have great significance, not only on academia, but also on the society as a whole.

The Special Session on Data Marketing as Cross-disciplinary Data Exchange and Collaboration (CDEC2020)

Chairs: Teruaki Hayashi, Yoshiaki Fukami, Didier Navez, Yukio Ohsawa

Time	Title	Presenter/Author
9:00-9:05	Opening Remarks	Teruaki Hayashi
9:05-9:30	Verification of Data Similarity using Metadata on a Data Exchange Platform (SP01205)	Hiroki Sakaji
9:30-9:55	Data Requests and Scenarios for Data Design of Unobserved Events in Corona-related Confusion Using TEEDA (SP01202)	Teruaki Hayashi
9:55-10:05	Coffee Break	
10:05-10:30	User authentication based on smartphone application usage patterns through learning classifier systems (SP01203)	Mhd Irvan
10:30-10:55	Hierarchical Graph Convolutional Network for Data Evaluation of Dynamic Graphs (SP01204)	Bin Wang
10:55-11:45	Invited Talk: TBA	TBA
11:45-11:55	General Comment	Yukio Ohsawa
11:55-12:00	Closing Remarks	

2nd Special Session on Machine Learning on Big Data (MLBD 2020)

Program – 11-12 December 2020

Virtual Event (EST Time)

The Special Session “Machine Learning on Big Data” (MLBD 2020) of the 2020 IEEE International Conference on Big Data (IEEE BigData 2020) follows the great success of three previous editions co-located with the IEEE BigData and IEEE ICMLA conference series and focuses on machine learning models, techniques and algorithms related to Big Data, a vibrant and challenging research context playing a leading role in the Machine Learning and Data Mining research communities. Big data is gaining attention from researchers, being driven among others by technological innovations (such as cloud interfaces) and novel paradigms (such as social networks). Devising and developing machine learning models, techniques and algorithms for big data represent a fundamental problem stirred-up by the tremendous range of critical applications incorporating machine learning tools in their core platforms.

For example, in application settings where big data arise and machine is useful, we recognize, among other things: (i) machine-learning-based processing (e.g., acquisition, knowledge discovery, and so forth) over large-scale sensor networks introduces important advantages over classical datamanagement- based approaches; similarly, (ii) medical and e-health information systems usually include successful machine learning tools for processing and mining very large graphs modelling patient-to-disease, patient-to-doctor, and patient-to-therapy networks; (iii) genome data management and mining can gain important benefits from machine learning algorithms. Some hot topics in machine learning on big data include: (i) machine learning on unconventional big data sources (e.g., large-scale graphs in scientific applications, stronglyunstructured social networks, and so forth); (ii) machine learning over massive big data in distributed settings; (iii) scalable machine learning algorithms; (iv) deep learning – models, principles, issues; (v) machine-learning-based predictive approaches; (vi) machine-learningbased big data analytics; (vii) privacy-preserving machine learning on big data; (viii) temporal analysis and spatial analysis on big data; (ix) heterogeneous machine learning on big data; (x) novel applications of machine learning on big data (e.g., healthcare, cybersecurity, smart cities, and so forth).

MLBD 2020 Detailed Program Schedule

Friday 11 December, 2020

Time	Title	Presenter/Author
08:55am – 09:00am	Session MLBD20_1: Session Opening Chair: Alfredo Cuzzocrea	
09:00am – 10:15am	Session MLBD20_2: Big Data Classification and Learning Chair: TBA	
09:00am – 09:15am	Taxonomic Classification of Objects with Convolutional Neural Networks	Sungryeol Yang, Geoffrey Fox, and Bokyoona Na
09:15am – 09:30am	Learning with Missing Data	Carlos Escobar, Jorge Arinez, Daniela Macias, and Ruben Morales-Menendez
09:30am – 09:45am	Zero-Shot Machine Learning Technique for Classification of Multi-User Big Data Workloads	Mikhail Genkin
09:45am – 10:00am	Extendable and invertible manifold learning with geometry regularized autoencoders	Andres Duque, Sacha Morin, Guy Wolf, and Kevin Moon
10:00am – 10:15am	Hyper-parameter optimization with REINFORCE and Masked Attention Auto-regressive Density Estimators	Chepurishri Krishna, Ashish Gupta, Swarnim Narayan, Himanshu Rai, and Diksha Manchanda
10:15am – 11:30am	Session MLBD20_3: Advanced Machine Learning Techniques over Big Data I Chair: TBA	
10:15am – 10:30am	Causal Inference with Correlation Alignment	Umar Isyaku Abdullahi, Spyros Samothrakis, and Maria Fasli
10:30am – 10:45am	Approximate Nearest-Neighbour Fields via Massively-Parallel Propagation-Assisted K-D Trees	Cosmin Eugen Oancea, Ties Robroek, and Fabian Gieseke
10:45am – 11:00am	Machine learning and OLAP on big COVID-19 data	Carson K. Leung, Yubo Chen, Calvin S.H. Hoi, Siyuan Shang, and Alfredo Cuzzocrea
11:00am – 11:15am	Identification and Prediction of Emerging Topics through Their Relationships to Existing Topics	Sukhwan Jung, Rituparna Datta, and Aviv Segev
11:15am – 11:30am	Semi-Unsupervised Learning: Clustering and Classifying using Ultra-Sparse Labels	Matthew Willetts, Stephen Roberts, and Chris Holmes
11:30am – 12:45pm	Session MLBD20_4: Intelligent Machine Learning Methods and Techniques over Big Data Chair: TBA	

11:30am – 11:45am	Data Reduction and Deep-Learning Based Recovery for Geospatial Visualization and Satellite Imagery	Jarin Tasnim and Debajyoti Mondal
11:45am – 12:00pm	Real-time Machine Learning Based on Hoeffding Decision Trees for Jamming Detection in 5G New Radio	Youness Arjouni and Saleh Faruque
12:00pm – 12:15pm	Forecasting People’s Action via Social Media Data	Lingfeng Shen, Zhuoming Liu, and Xiongtao Zhou
12:15pm – 12:30pm	A Novel Deep Learning Character-Level Solution to Detect Language and Printing Style from a Bilingual Scanned Document	AKM Shahariar Azad Rabby, Md. Majedul Islam, Nazmul Hasan, Jebun Nahar, and Fuad Rahman
12:30pm – 12:45pm	Sampling Approach Matters: Active Learning for Robotic Language Acquisition	Nisha Pillai, Edward Raff, Francis Ferraro, and Cynthia Matuszek
12:45am – 02:15pm	Session MLBD20_5: Architectures and Techniques for Advanced Machine Learning over Big Data Chair: TBA	
12:45am – 01:00pm	A Sensitivity Analysis (and Practitioners’ Guide to) of DeepSORT for Low Frame Rate Video	Manikandan Ravikiran, Yuichi Nonaka, and Nestor Mariyasagayam
01:00pm – 01:15pm	Adaptive and Efficient Streaming Time Series Forecasting with Lambda Architecture and Spark	Arjun Pandya, Oluwatobiloba Odunsi, Chen Liu, Alfredo Cuzzocrea, and Jianwu Wang
01:15pm – 01:30pm	A Streaming Machine Learning Framework for Online Aggression Detection on Twitter	Herodotos Herodotou, Despoina Chatzakou, and Nicolas Kourtellis
01:30pm – 01:45pm	Effects of Missing Members on Classifier Ensemble Accuracy	Alec Austin and Ryan Benton
01:45pm – 02:00pm	Community Detection using Semi-supervised Learning with Graph Convolutional Network on GPUs	Naw Safrin Sattar and Shaikh Arifuzzaman
Saturday 12 December, 2020		
Time	Title	Presenter/Author
09:00am – 10:15am	Session MLBD20_6: Foundations of Machine Learning over Big Data Chair: TBA	
09:00am – 09:15am	Large-scale Sparse Structural Node Representation	Edoardo Serra, Mikel Joaristi, and Alfredo Cuzzocrea
09:15am – 09:30am	Feature-based Distant Domain Transfer Learning	Shuteng Niu, Yihao Hu, Jian Wang, Yongxin Liu, and Houbing Song
09:30am – 09:45am	Exploiting Data Entropy for Neural Network Compression	Tse-Wen Chen, Pangfeng Liu, and Jan-Jan Wu
09:45am – 10:00am	A Deep Reinforcement Learning Framework for Instructional Sequencing	YanJun Pu, Caimeng Wang, and Wenjun Wu
10:00am – 10:15am	Optimizing Stochastic Gradient Descent Using the Angle Between Gradients	Chongya Song, Alexander Pons, and Kang Yen
10:15am – 11:30am	Session MLBD20_7: Advanced Machine Learning Techniques over Big Data II Chair: TBA	
10:15am – 10:30am	An Autocorrelation-based LSTM-Autoencoder for Anomaly Detection on Time-Series Data	Hajar Homayouni, Sudipto Ghosh, Indrakshi Ray, Shlok Gondalia, Jerry Duggan, and Michael G. Kahn
10:30am – 10:45am	Weight Prediction for Variants of Weighted Directed Networks	Dong Quan Nguyen, Lin Xing, and Lizhen Lin
10:45am – 11:00am	Bag of Symbols for Time Series Distance Measurement and Applications	Pejman Khadivi
11:00am – 11:15am	ONet – A Temporal Meta Embedding Network for MOOC Dropout Prediction	Siby Charley Pulikottil and Manish Gupta
11:15am – 11:30am	Understanding and Mitigating Threats from Android Hybrid Apps Using Machine Learning	AK Singh and Navneet Goyal
11:30am – 12:45pm	Session MLBD20_8: Big Data Clustering and Forecasting Chair: TBA	
11:30am – 11:45am	Pseudo-Reconfigurable Heterogeneous Solution for Accelerating Spectral Clustering	Mihaela Malita, George-Vladut Popescu, and Gheorghe Stefan

11:45am – 12:00pm	Performance Evaluation of Tree-based Models for Big Data Load Forecasting using Randomized Hyper parameter Tuning	Ameema Zainab, Ali Ghrayeb, Mahdi Houchati, Shady Khalil, and Haitham Abu-Rub
12:00pm – 12:15pm	Dynamic Graph Neural Network for Traffic Forecasting in Wide Area Networks	Tanwi Mallick, Mariam Kiran, Bashir Mohammed, and Prasanna Balaprakash
12:15pm – 12:30pm	FedCluster: Boosting the Convergence of Federated Learning via Cluster-Cycling	Cheng Chen, Ziyi Chen, Yi Zhou, and Bhavya Kailkhura
12:30pm – 12:45pm	Leveraging Natural Language Processing to Understand Public Outlook Towards the Influenza Vaccination	Ankita Agarwal, William L. Romine, and Tanvi Banerjee
12:45am – 01:30pm	Session MLBD20_9: Machine Learning Approaches for Recommendations over Big Data Chair: TBA	
12:45am – 01:00pm	Addressing Cold Start in Recommender Systems with Hierarchical Graph Neural Networks	Ivan Maksimov, Rodrigo Rivera-Castro, and Evgeny Burnaev
01:00pm – 01:15pm	Session-based Recommendation model based on Multiple Neural Networks hybrid extraction feature	Huaxiong Yao, Jiabei Hu, Wenqi Xie, Wei Xie, and Yang Huang
01:15pm – 01:30pm	Recommendations of Compatible Accessories in e-Commerce	San He Wu, Unaiza Ahsan, Mingming Guo, Simon Hughes, Xiquan Cui, and Khalifeh Al Jadda
01:30pm – 02:00pm	Session MLBD20_10: Miscellaneous Chair: TBA	
01:30pm – 01:45pm	An Innovative Framework for Supporting Remote Sensing in Image Processing Systems via Deep Transfer Learning	Oxana Korzh, Ashish Sharma, Mikel Joaristi, Edoardo Serra, and Alfredo Cuzzocrea
01:45pm – 02:00pm	Virtual Adversarial Active Learning	Chin-Feng Yu and Hsing-Kuo Pao
02:00pm – 02:15pm	Session MLBD20_11: Session Closing Chair: Alfredo Cuzzocrea	

BigData Cup Challenges

12/10(Thu)9:00- USA Eastern Standard Time

12/10(Thu)23:00- JST, 12/10(Thu)20:30- IST

11 mins presentation 4 mins Q&A

BigData Cup: Global Road Damage Detection Challenge 2020		
Chairs: Hiroya Maeda <maedahi@iis.u-tokyo.ac.jp>		
Time	Title	Presenter/Author
9:00-9:05	Opening Remarks	
9:05-9:20	Global Road Damage Detection: State-of-the-art Solutions	Deeksha Arya, Hiroya Maeda, Sanjay Kumar Ghosh, Durga Toshniwal, Hiroshi Omata, and Yoshihide Sekimoto
9:20-9:35	An Efficient and Scalable Deep Learning Approach for Road Damage Detection	Sadra Naddaf Sargh, Mohammad Mahdi Naddaf Sargh, Amir Reza Kashani, and Hassan Zargarzadeh
9:35-9:50	CFM: A Consistency Filtering Mechanism for Road Damage Detection	Zixiang Pei, Xiubao Zhang, Rongheng Lin, Haifeng Shen, Jian Tang, and Yi Yang
9:50-10:00	Road Damage Detection using Deep Ensemble Learning	Keval Doshi and Yasin Yilmaz
	Coffee Break	
10:15-10:30	Exploring the Tricks for Road Damage Detection with A One-Stage Detector	Xiaoguang Zhang, Xuan Xia, Nan Li, Lin Ma, Junlin Song, and Ning Ding
10:30-10:45	Detecting Various Road Damage Types in Global Countries Utilizing Faster R-CNN	Felix Kortmann, Kevin Talits, Pascal Fassmeyer, Alexander Warnecke, Nicolas Meier, Jens Heger, Paul Drews, and Burkhardt Funk
10:45-11:00	Road Damage Detection Using YOLO with Smartphone Images	Dongjun Jeong
11:00-11:15	Road Damage Detection and Classification with Detectron2 and Faster R-CNN	Vung Pham, Chau Pham, and Tommy Dang
	Coffee Break	
11:30-11:45	Ensemble Learning for Road Damage Detection and Classification	Vinuta Hedge, Dweep Trivedi, Abdullah Alfarrarjeh, Aditi Deepak, Seon Ho Kim, and Cyrus Shahabi
11:45-12:00	Deep Network For Road Damage Detection	Yuming Liu, Xiaoyong Zhang, Bingzhen Zhang, and Zhenwu Chen
12:00-12:15	CNN Model & Tuning for Global Road Damage Detection	Rahul Vishwakarma and Ravigopal Vennelakanti
12:15-12:30	FasterRCNN Monitoring of Road Damages: Competition and Deployment	Tristan Hascoet, Yihao Zhang, Andreas Persch, Ryoichi Takashima, Tetsuya Takiguchi, and Yasuo Arikawa
12:30-12:45	Deep Learning Frameworks for Pavement Distress Classification: A Comparative Analysis	Vishal Mandal, Abdul Rashid Mussah, and Yaw Adu-Gyamfi
12:45-13:00	Closing Remarks & Awards Ceremony	

Posters

Poster ID	Accept Posters
P201	Abdullah Almoqbil, Brian O'Connor, Rich Anderson, and Jibril Shittu, <i>The correlation between substance abuse and crime in the United States</i> Author Email(s): abdullah.almoqbil@untsystem.edu, brian.oconnor@unt.edu, rich.anderson@untsystem.edu, JibrilShittu@my.unt.edu Contact Person: Abdullah Almoqbil <abdullah.almoqbil@untsystem.edu>
P203	Mark Hamilton, Anand Raman, Christina Lee, Lucy Zhang, Lei Zhang, and William Freeman, <i>Large-Scale Intelligent Microservices</i> Author Email(s): marhamil@microsoft.com, aram@microsoft.com, chril@microsoft.com, zhangly@mit.edu, leizhang@microsoft.com, billf@mit.edu Contact Person: Mark Hamilton <marhamil@microsoft.com>
P204	Manoj M, <i>Integrating Polystore RDBMS with Common In-Memory Data</i> Author Email(s): mm42526w@pace.edu Contact Person: Manoj M <mm42526w@pace.edu>
P205	Manoj M, <i>Approximate Query Processing for Big Data in Heterogeneous Databases</i> Author Email(s): mm42526w@pace.edu Contact Person: Manoj M <mm42526w@pace.edu>
P206	Byron Gao, <i>Pattern Exploration as Keyword Search</i> Author Email(s): bgao@txstate.edu Contact Person: Byron Gao <bgao@txstate.edu>
P207	Byron Gao and Alexander Katrompas, <i>A Preliminary Experimental Analysis on RateMyProfessors</i> Author Email(s): bgao@txstate.edu, amk181@txstate.edu Contact Person: Byron Gao <bgao@txstate.edu>
P208	Manoj M, <i>Survey of the use of digital technologies to combat COVID-19</i> Author Email(s): mm42526w@pace.edu Contact Person: Manoj M <mm42526w@pace.edu>
P209	Michael Shur, <i>Describing and Predicting COVID19 Evolution Using Pandemic Equation</i> Author Email(s): shurm@rpi.edu Contact Person: Michael Shur <shurm@rpi.edu>
P210	Xu Du, Matthew Kowalski, Aparna Varde, and Boxiang Dong, <i>LSOMP: Large Scale Ordinance Mining Portal</i> Author Email(s): dux3@montclair.edu, kowalskim6@montclair.edu, vardea@montclair.edu, dongb@montclair.edu Contact Person: Xu Du <dux3@montclair.edu>
P211	Ranjeet Devarakonda, Jitendra Kumar, and Giri Prakash, <i>Clustering-Based Predictive Analytics to Improve Scientific Data Discovery</i> Author Email(s): devarakondar@ornl.gov Contact Person: Ranjeet Devarakonda <devarakondar@ornl.gov>
P212	Jianchuan Li, Peiquan Jin, and Shouhong Wan, <i>Adaptive Lazy Compaction with High Stability and Low Latency for Data-Intensive Systems</i> Author Email(s): jpq@ustc.edu.cn, lij@ustc.edu.cn, wansh@ustc.edu.cn Contact Person: Peiquan Jin <jpq@ustc.edu.cn>
P213	Nicolas Schroeder, <i>Predicting the spread of Covid-19 through quantified cultural dimensions</i> Author Email(s): nisc19ad@student.cbs.dk Contact Person: Nicolas Schroeder <nisc19ad@student.cbs.dk>
P216	Eric Bax and Charlotte Bax, <i>SAFE – Secure Aggregated Frequency Estimates</i> Author Email(s): ebax@verizonmedia.com, 23baxc@flintridgeprep.org Contact Person: Eric Bax <ebax@verizonmedia.com>
P217	Jia Xie, Shouhong Wan, and Peiquan Jin, <i>Fast and Effective Object Classification for Big Image Data</i> Author Email(s): jpq@ustc.edu.cn, geroi@mail.ustc.edu.cn, wansh@ustc.edu.cn Contact Person: Peiquan Jin <jpq@ustc.edu.cn>

P218	Oleksandr Letychevskiy and Yaroslav Hryniuk, <i>Machine Learning Methods for Improving Vulnerability Detection in Low-level Code</i> Author Email(s): oleksandr.letychevskiy@litsoft.com.ua, yaroslav.hryniuk@gmail.com Contact Person: Oleksandr Letychevskiy <oleksandr.letychevskiy@litsoft.com.ua>
P219	Chih-Chieh Yang, Giacomo Domeniconi, Leili Zhang, and Guojing Cong, <i>Design of AI-Enhanced Drug Lead Optimization Workflow for HPC and Cloud</i> Author Email(s): chih.chieh.yang@ibm.com, giacomo.domeniconi1@ibm.com, zhangle@us.ibm.com, gcong@us.ibm.com Contact Person: Chih-Chieh Yang <chih.chieh.yang@ibm.com>
P220	Yi Wang, Peiquan Jin, and Shouhong Wan, <i>HotKey-LSM: A Hotness-Aware LSM-Tree for Big Data Storage</i> Author Email(s): jpq@ustc.edu.cn, wykde@mail.ustc.edu.cn, wansh@ustc.edu.cn Contact Person: Peiquan Jin <jpq@ustc.edu.cn>
P221	Vinu ELLAMPALLIL VENUGOPAL and Martin THEOBALD, <i>Effective Stream Data Processing using Asynchronous Iterative Routing Protocol</i> Author Email(s): vinu.venugopal@uni.lu, martin.theobald@uni.lu Contact Person: Vinu ELLAMPALLIL VENUGOPAL <vinu.venugopal@uni.lu>
P222	Jerome Heng, Junhua Liu, and Kwan Hui Lim, <i>Urban Crowdsensing using Social Media: An Empirical Study on Transformer and Recurrent Neural Networks</i> Author Email(s): kwanhui_lim@sutd.edu.sg, jerome_heng@mymail.sutd.edu.sg, junhua_liu@mymail.sutd.edu.sg Contact Person: Kwan Hui Lim <kwanhui_lim@sutd.edu.sg>
P223	He Li, Shiyu Zhang, Liangcai Su, Hongjie Huang, Duo Jin, and Xuejiao Li, <i>GraphSAnet: A Graph Neural Network and Self Attention Based Approach for Spatial Temporal Prediction in Sensor Network</i> Author Email(s): suliangcai@stu.xidian.edu.cn, heli@xidian.edu.cn, sy_zhang@stu.xidian.edu.cn, huanghongjie@stu.xidian.edu.cn, djin@stu.xidian.edu.cn, lixuejiao@stu.xidian.edu.cn Contact Person: Liangcai Su <suliangcai@stu.xidian.edu.cn>
P224	Xuejiao Tang, Jiong Qiu, Ruijun Chen, Wenbin Zhang, Vasileios Iosifidis, Zhen Liu, Wei Meng, Mingli Zhang, and Ji Zhang, <i>A Data-driven Human Responsibility Management System</i> Author Email(s): colin@hotmail.com, xuejiao.tang@stud.uni-hannover.de, n78083016@mail.ncku.edu.tw, wenbinzhang@umbc.edu, iosifidis@l3s.de, liu.zhen@gdpu.edu.cn, mnancy@bjfu.edu.cn, mingli.zhang@mcgill.ca Contact Person: Jiong Qiu <colin@hotmail.com>
P225	Corey Hannum, Rui Li, and Weitian Wang, <i>Trust or Not?: A Computational Robot-Trusting-Human Model for Human-Robot Collaborative Tasks</i> Author Email(s): wangw@montclair.edu, hannumc1@montclair.edu, ruil@clemson.edu Contact Person: Weitian Wang <wangw@montclair.edu>
P226	Odysseas Tsilingiridis and Alexandros Karakasidis, <i>MILMS: A Microservices-based Learning Management System</i> Author Email(s): std115787@ac.eap.gr, a.karakasidis@uom.edu.gr Contact Person: Odysseas Tsilingiridis <std115787@ac.eap.gr>
P227	Andreas Messalas, Christos Aridas, and Yannis Kanellopoulos, <i>Evaluating MASHAP as a faster alternative to LIME for model-agnostic machine learning interpretability</i> Author Email(s): andreas@code4thought.eu, chris@code4thought.eu, yannis@code4thought.eu Contact Person: Andreas Messalas <andreas@code4thought.eu>
P228	Christopher Conti, Aparna Varde, and Weitian Wang, <i>Task quality optimization in collaborative robotics</i> Author Email(s): contic5@montclair.edu, vardea@mail.montclair.edu, wangw@montclair.edu Contact Person: Christopher Conti <contic5@montclair.edu>
P229	Aiswarya Sriram, Advithi Nair, Alka Simon, Subramaniam Kalambur, and Dinkar Sitaram, <i>A Study on the Causes of Garbage Collection in Java for Big Data Workloads</i> Author Email(s): aiswarya.spaa@gmail.com, advithi.nair@gmail.com, alkasimon23@gmail.com, subramaniamkv@pes.edu, dinkar.sitaram@gmail.com Contact Person: Aiswarya Sriram <aiswarya.spaa@gmail.com>
P230	Julia Mell and Anuar Imanbayev, <i>Impact of BGM Point Profiles on Glycemic Variability Prediction Accuracies</i> Author Email(s): ajmell7@gmail.com, anui@novonordisk.com Contact Person: Julia Mell <ajmell7@gmail.com>
P231	Yihyun Nam and Sangwhan Cha, <i>Correlation Analysis between Median Income level of District and Quality of Medical Service in Seoul, Korea</i> Author Email(s): iindes@daum.net, scha@harrisburgu.edu Contact Person: Yihyun Nam <iindes@daum.net>
P232	Fan MO, Huida JIAO, Shun MORISAWA, Makoto NAKAMURA, Koichi KIMURA, Hisanori FUJISAWA, Masafumi OHTSUKA, and Hayato YAMANA, <i>Real-Time Periodic Advertisement Recommendation Optimization using Ising Machine</i> Author Email(s): bakubonn@toki.waseda.jp, jjiao@yama.info.waseda.ac.jp, hiroshun@yama.info.waseda.ac.jp, marumakan@fujitsu.com, k.kimura@fujitsu.com, h.fujisawa@fujitsu.com, m.ohtsuka@geniee.co.jp,

	yamana@yama.info.waseda.ac.jp Contact Person: Fan MO <bakubonn@toki.waseda.jp>
P233	Boyu Liu, Duanyue Yun, Xin Guo, Xiao Ji, Huiyu Song, Shirish Singh, and Gail Kaiser, <i>Detecting Sensor-Based Repackaged Malware</i> Author Email(s): shirish@cs.columbia.edu, bl2791@columbia.edu, dy2400@columbia.edu, xg2331@columbia.edu, xj2247@columbia.edu, hs3160@columbia.edu, kaiser@cs.columbia.edu Contact Person: Shirish Singh <shirish@cs.columbia.edu>
P234	Moonhyung Lee and Sangwhan Cha, <i>Towards Personalized hybrid Recommender System using Average Time Intervals</i> Author Email(s): m4a1bruce@gmail.com, scha@harrisburgu.edu Contact Person: Moonhyung Lee <m4a1bruce@gmail.com>
P235	Ruohan Zong, <i>On Privileged Information Driven Robust Face Verification: A Siamese Convolutional Neural Network Approach</i> Author Email(s): rz2538@columbia.edu Contact Person: Ruohan Zong <rz2538@columbia.edu>
P236	Koji Iwanuma, Kento Yajima, and Yoshitaka Yamamoto, <i>Mining Consistent, Non-Redundant and Minimal Negative Rules Based on Minimal Generators</i> Author Email(s): iwanuma@yamanashi.ac.jp Contact Person: Koji Iwanuma <iwanuma@yamanashi.ac.jp>
P237	Qingling Feng, Ruisheng Shi, Qifeng Luo, Lina Lan, and Jinqiao Shi, <i>Scalable Blind Matching: An Efficient Ciphertext Matching Scheme for Content-Based Pub/Sub Cloud Services</i> Author Email(s): fqLing@bupt.edu.cn, shiruisheng@bupt.edu.cn, luofeng@bupt.edu.cn, lanlina@bupt.edu.cn, shijinqiao@bupt.edu.cn Contact Person: Qingling Feng <fqLing@bupt.edu.cn>
P238	Zeno Bardelli, Ines Arous, Philippe Cudré-Mauroux, and Ljiljana Dolamic, <i>SwissFinder: Identifying Swiss Websites from Unstructured Content</i> Author Email(s): ines@exascale.info, zeno.bardelli@unifr.ch, pcm@unifr.ch, Ljiljana.Dolamic@armasuisse.ch Contact Person: Ines Arous <ines@exascale.info>
P239	Junchen Zhang, Yongping Luo, Peiquan Jin, and Shouhong Wan, <i>Optimizing Adaptive Radix Trees for NVM-Based Hybrid Memory Architecture</i> Author Email(s): jpq@ustc.edu.cn, jc Zhang2@mail.ustc.edu.cn, ypluo@mail.ustc.edu.cn, wansh@ustc.edu.cn Contact Person: Peiquan Jin <jpq@ustc.edu.cn>
P240	Evangelia Kavakli, Rizos Sakellariou, Iliada Eleftheriou, and Julien Etienne Mascolo, <i>Towards a Multi-Perspective Methodology for Big Data Requirements</i> Author Email(s): evangelia.kavakli-2@manchester.ac.uk, rizos@manchester.ac.uk, iliada.eleftheriou@manchester.ac.uk, julienetienne.mascolo@crf.it Contact Person: Evangelia Kavakli <evangelia.kavakli-2@manchester.ac.uk>
P241	Le Phan, Thanh-Nam Doan, and Mina Sartipi, <i>Understanding the Effect of COVID-19 on Fuel Consumption of Public Transportation: The case study of Chattanooga, TN</i> Author Email(s): bbz181@mocs.utc.edu, thanh-nam-doan@utc.edu, mina-sartipi@utc.edu Contact Person: Le Phan <bbz181@mocs.utc.edu>
P242	Qiqiang Xu, Ji Zhang, Youwen Zhu, Bohan Li, Donghai Guan, and Xin Wang, <i>A Block-Level RNN Model for Resume Block Classification</i> Author Email(s): Ji.Zhang@usq.edu.au, 1536801671@qq.com, zhuyw@nuaa.edu.cn, bhli@nuaa.edu.cn, dhguan@nuaa.edu.cn, xinwang.ed@gmail.com Contact Person: Ji Zhang <Ji.Zhang@usq.edu.au>
P243	Danielle Lambion, Michael Josten, Femi Olumofin, and Martine De Cock, <i>Malicious DNS Tunneling Detection in Real-Traffic DNS Data</i> Author Email(s): mjosten@uw.edu, dlambion@uw.edu, folumofin@infoblox.com, mdcock@uw.edu Contact Person: Michael Josten <mjosten@uw.edu>
P244	Jun Hao and Xinghui Zhao, <i>An End-to-End Toolkit for Social Network Analysis</i> Author Email(s): x.zhao@wsu.edu, jun.hao@wsu.edu Contact Person: Xinghui Zhao <x.zhao@wsu.edu>
P245	Minhyeok Kweun, Woo-Yeon Lee, Goeun Kim, Jisoo Hwang, and Yoonkyong Lee, <i>Lineage Checkpoint Approach for Long-lineage Problem in Apache Spark</i> Author Email(s): mh.kweun@samsung.com, wooyeon0.lee@samsung.com, ge326.kim@samsung.com, jisoo0.hwang@samsung.com, yk6629.lee@samsung.com Contact Person: Minhyeok Kweun <mh.kweun@samsung.com>
P246	Hyeon-Jun Jang, Yin-Goo Yim, and Hyun-Wook Jin, <i>Vertical Autoscaling of GPU Resources for Machine Learning in the Cloud</i>

	<p>Author Email(s): jinh@konkuk.ac.kr, guswns531@konkuk.ac.kr, ygyim@konkuk.ac.kr Contact Person: Hyun-Wook Jin <jinh@konkuk.ac.kr></p>
P247	<p>Ravid Shwartz-Ziv, Amitai Armon, and Itamar Ben Ari, <i>Spatial-Temporal Convolutional Network for Spread Prediction of COVID-19</i> Author Email(s): ravid.ziv@intel.com, amitai.armon@intel.com, itamar.ben-ari@intel.com Contact Person: Ravid Shwartz-Ziv <ravid.ziv@intel.com></p>
P248	<p>Hani Ramadhan and Joonho Kwon, <i>Learning Minimum Bounding Rectangles for Efficient Trajectory Similarity Search</i> Author Email(s): hani042@pusan.ac.kr, jhkwon@pusan.ac.kr Contact Person: Hani Ramadhan <hani042@pusan.ac.kr></p>
P249	<p>Qiqiang Xu, Ji Zhang, Zenghui Xu, Yonglong Luo, Fulong Chen, Xiaoyao Zheng, and Gaoming Yang, <i>Effective Tuple-based Anonymization for Massive Streaming Categorical Data</i> Author Email(s): zhangji77@gmail.com, 1536801671@qq.com, xuzenghui@zhejianglab.com, ylluo@ustc.edu.cn, long005@mail.ahnu.edu.cn, zxiaoyao_2000@163.com, ygm868@163.com Contact Person: Ji Zhang <zhangji77@gmail.com></p>
P250	<p>Chuen-Min Huang, Kuo-Lin Lu, Yi-Ying Cheng, and Yu-Chen Peng, <i>Generating Chinese Classical Poetry with Quatrain Generation Model (QGM) Using Encoder-Decoder LSTM</i> Author Email(s): jennyhuang921@gmail.com, bullet.kids@gmail.com, alicelisecheng@gmail.com, peng860812@gmail.com Contact Person: Chuen-Min Huang <jennyhuang921@gmail.com></p>
P251	<p>Hoang H. Nguyen, Sergej Zerr, and Tuan-Anh Hoang, <i>On Node Embedding of Uncertain Networks</i> Author Email(s): ehong@l3s.de, szerr@l3s.de, hoangtuananh@hus.edu.vn Contact Person: Hoang H. Nguyen <ehong@l3s.de></p>
P252	<p>Chandan Misra, Utkarsh Parasarampuria, Sourangshu Bhattacharya, and Soumya K. Ghosh, <i>On Distributed Solution for Simultaneous Linear Symmetric Systems</i> Author Email(s): chandan@xub.edu.in, utkarsh@rawky-studios.com, sourangshu@cse.iitkgp.ac.in, skg@cse.iitkgp.ac.in Contact Person: Chandan Misra <chandan@xub.edu.in></p>
P253	<p>Utkarsh Parasarampuria, Chandan Misra, and Sourangshu Bhattacharya, <i>An Optimized Distributed Recursive Matrix Multiplication for Arbitrary Sized Matrices</i> Author Email(s): chandan@xub.edu.in, utkarsh@rawky-studios.com, sourangshu@cse.iitkgp.ac.in Contact Person: Chandan Misra <chandan@xub.edu.in></p>
P254	<p>Hidehiro Matsumoto and Akira Ishii, <i>An Analysis Approach of Messaging Mechanism on Social Networking Services</i> Author Email(s): hmatsumoto@nvd.co.jp, ishii@tottori-u.ac.jp Contact Person: Hidehiro Matsumoto <hmatsumoto@nvd.co.jp></p>
P255	<p>Mehdi Dadfarnia, <i>Risk-Based Evaluation of Diagnostic Algorithms for Multi-stage Manufacturing</i> Author Email(s): mahdid@gmail.com Contact Person: Mehdi Dadfarnia <mahdid@gmail.com></p>
P256	<p>Srinivas Danda, Ji Zhang, Xiaohui Tao, Jerry Chun-Wei Lin, and Wenbin Zhang, <i>Context-aware Adaptive Outlier Detection in Trajectory Data</i> Author Email(s): zhangji77@gmail.com, rao.danda@gmail.com, Xiaohui.Tao@usq.edu.au, jerrylin@ieee.org, wenbinzhang@umbc.edu Contact Person: Ji Zhang <zhangji77@gmail.com></p>
P257	<p>Julius Roeder, Benjamin Rouxel, and Clemens Grelck, <i>Q-learning for Statically Scheduling DAGs</i> Author Email(s): j.roeder@uva.nl, b.rouxel@uva.nl, c.grelck@uva.nl Contact Person: Julius Roeder <j.roeder@uva.nl></p>
P258	<p>Sebastian Trinks and Carsten Felden, <i>A differentiation between Image Mining and Computer Vision in the application area of Big Data</i> Author Email(s): sebastian.trinks@bwl.tu-freiberg.de, carsten.felden@bwl.tu-freiberg.de Contact Person: Sebastian Trinks <sebastian.trinks@bwl.tu-freiberg.de></p>
P259	<p>Zishun Feng, Ming Tu, Rui Xia, Yuxuan Wang, and Ashok Krishnamurthy, <i>Self-Supervised Audio-Visual Representation Learning for in-the-wild Videos</i> Author Email(s): fzs@cs.unc.edu, ashok@renci.org Contact Person: Zishun Feng <fzs@cs.unc.edu></p>
P260	<p>Yasuko Kawahata, <i>Stochastic Process for Analyzing Speech on the Web with Consideration of Media Mediation in Large-scale Broadcast Events in Japan</i> Author Email(s): kawahata@rikkyo.ac.jp Contact Person: Yasuko Kawahata <kawahata@rikkyo.ac.jp></p>
P261	<p>Mark Cheung and Jose' M. F. Moura, <i>Graph Neural Networks for COVID-19 Drug Discovery</i> Author Email(s): markcheu@andrew.cmu.edu, moura@andrew.cmu.edu Contact Person: Mark Cheung <markcheu@andrew.cmu.edu></p>

P262	<p>Simone Monaco, Andrea Pasini, Daniele Apiletti, Luca Colomba, Paolo Garza, and Elena Baralis, <i>Improving Wildfire Severity Classification of Deep Learning U-Nets from Satellite Images</i></p> <p>Author Email(s): danielle.apiletti@polito.it, simone.monaco@studenti.polito.it, andrea.pasini@polito.it, luca.colomba@polito.it, paolo.garza@polito.it, elena.baralis@polito.it</p> <p>Contact Person: Daniele Apiletti <danielle.apiletti@polito.it></p>
P263	<p>William Ledbetter and John Springer, <i>Help Speech Identification using HathiTrust and ABBYY Fine Reader</i></p> <p>Author Email(s): wledbett@purdue.edu, jaspring@purdue.edu</p> <p>Contact Person: William Ledbetter <wledbett@purdue.edu></p>
P264	<p>Colin Berry, <i>The Diffusion of Information: The Impact of Sentiment and Topic on Retweets</i></p> <p>Author Email(s): berrycolin17@gmail.com</p> <p>Contact Person: Colin Berry <berrycolin17@gmail.com></p>
P265	<p>Emad Felemban, Faizan Ur Rehman, Ahmad Muaz Qamar, and Akhlaq Ahmed, <i>A Framework to Extract Trip Informatics for Individuals and Groups Using Mobile Sensory Data in Mass Gathering Events</i></p> <p>Author Email(s): faizanurrehman@gmail.com, eafelemban@uqu.edu.sa, amqamar@uqu.edu.sa, professorakhlaq@gmail.com</p> <p>Contact Person: Faizan Ur Rehman <faizanurrehman@gmail.com></p>
P266	<p>Anna Pyayt, Rituparna Khan, Robert Brzozowski, Prahathes Eswara, and Michael Gubanov, <i>Rapid Antibiotic Susceptibility Analysis via Microscopy and Machine Learning</i></p> <p>Author Email(s): gubanov@cs.fsu.edu, anna.pyayt@gmail.com, rtp@gmail.com, rb@usf.edu, pess@usf.edu</p> <p>Contact Person: Michael Gubanov <gubanov@cs.fsu.edu></p>
P267	<p>Justine Caylor, Somiya Metu, and Adrienne Raglin, <i>The Role of Multi-Criteria Decision-Making in a Sentry Agents Framework Utilizing Uncertainty of Information</i></p> <p>Author Email(s): justine.p.caylor.ctr@mail.mil, somiya.metu.civ@mail.mil, adrienne.raglin2.civ@mail.mil</p> <p>Contact Person: Justine Caylor <justine.p.caylor.ctr@mail.mil></p>
P268	<p>Drake Johnson, Keith Register, Brian Davison, and Jeff Heflin, <i>An Exploratory Interface for Dataset Repositories Using Cell-Centric Indexing</i></p> <p>Author Email(s): joh0057@calu.edu, kregister26@gmail.com, davison@cse.lehigh.edu, heflin@cse.lehigh.edu</p> <p>Contact Person: Drake Johnson <joh0057@calu.edu></p>
P269	<p>Rituparna Khan and Michael Gubanov, <i>Towards Tabular Embeddings, Training the Relational Models</i></p> <p>Author Email(s): gubanov@cs.fsu.edu, rk18g@my.fsu.edu</p> <p>Contact Person: Michael Gubanov <gubanov@cs.fsu.edu></p>
P270	<p>Rituparna Khan and Michael Gubanov, <i>WEBLENS: Towards Web-scale Data Integration, Training the Models</i></p> <p>Author Email(s): gubanov@cs.fsu.edu, rk18g@my.fsu.edu</p> <p>Contact Person: Michael Gubanov <gubanov@cs.fsu.edu></p>
P271	<p>Kavya Guntupally and Ranjeet Devarakonda, <i>Automated Indexing Of Structured Scientific Metadata Using Apache Solr</i></p> <p>Author Email(s): guntupallyk@ornl.gov, devarakondar@ornl.gov</p> <p>Contact Person: Kavya Guntupally <guntupallyk@ornl.gov></p>
P272	<p>Yichen Zhang and Thilanka Munasinghe, <i>Exploring Image Segmentation Techniques to Detect High Latitude Dust on Satellite Images</i></p> <p>Author Email(s): yichenzhangycz@gmail.com, Munast@rpi.edu</p> <p>Contact Person: Yichen Zhang <yichenzhangycz@gmail.com></p>
P273	<p>Joshua Chuah and Thilanka Munasinghe, <i>Exploratory Data Analysis to Understand Social Determinants Important to Global Neonatal Mortality Rate</i></p> <p>Author Email(s): chuahj2@rpi.edu, munast@rpi.edu</p> <p>Contact Person: Joshua Chuah <chuahj2@rpi.edu></p>
P275	<p>Arjumand Younus and Muhammad Atif Qureshi, <i>Combining BERT with Contextual Linguistic Features for Identification of Propaganda Spans in News Articles</i></p> <p>Author Email(s): arjumand.younus@ucd.ie, muhammadatif.qureshi@tudublin.ie</p> <p>Contact Person: Arjumand Younus <arjumand.younus@ucd.ie></p>
P276	<p>Maninder Singh and Vaibhav Anu, <i>Graph Based CIA in Requirements Engineering</i></p> <p>Author Email(s): anuv@montclair.edu, msingh@stcloudstate.edu</p> <p>Contact Person: Vaibhav Anu <anuv@montclair.edu></p>
P278	<p>Chau-Lin (Charly) Huang and Thilanka Munasinghe, <i>Exploring Various Applicable Techniques to Detect Smoke on the Satellite Images</i></p> <p>Author Email(s): huangc11@rpi.edu, munast@rpi.edu</p> <p>Contact Person: Chau-Lin (Charly) Huang <huangc11@rpi.edu></p>
P279	<p>Kaiming Fu, Yulu Jin, and Zhousheng Chen, <i>Test Set Optimization by Machine Learning Algorithms</i></p> <p>Author Email(s): kmfu@ucdavis.edu, yuljin@ucdavis.edu, zhche@ucdavis.edu</p> <p>Contact Person: Kaiming Fu <kmfu@ucdavis.edu></p>

P280	Xiantian Zhou and Carlos Ordonez, <i>Matrix Multiplication with SQL Queries for Graph Analytics</i> Author Email(s): xiantianzhou@gmail.com, carlos@central.uh.edu Contact Person: Xiantian Zhou <xiantianzhou@gmail.com>
P281	Catalina M Jaramillo, Paul Squires, Harold Kaufman, Andre Mendes da Silva, and Julian Togelius, <i>Word Embedding For Job Market Spatial Representation: Tracking Changes And Predicting Skills Demand.</i> Author Email(s): cmj383@nyu.edu, ps2937@nyu.edu, hgk224@nyu.edu, andre.mendes@nyu.edu, julian.togelius@nyu.edu Contact Person: Catalina M Jaramillo <cmj383@nyu.edu>
P282	Carlos Ordonez, Sikder Tahsin Al-Amin, and Ladjel Bellatreche, <i>An ER-Flow Diagram for Big Data</i> Author Email(s): stahsin.cse@gmail.com, carlos@central.uh.edu, ladjel.bellatreche@ensma.fr Contact Person: Sikder Tahsin Al-Amin <stahsin.cse@gmail.com>
P283	Carlos Garcia-Alvarado and Carlos Ordonez, <i>Querying Big Source Code</i> Author Email(s): carlosgarciaalvarado@gmail.com, carlos@central.uh.edu Contact Person: Carlos Garcia-Alvarado <carlosgarciaalvarado@gmail.com>
P284	Jacob Balma, Aaron Vose, Yuri Peterson, Amar Chittiboyina, Pankaj Pandey, Charles Yates, Ikhlous Khan, and Sreenivas Sukumar, <i>Deep Learning Predicts Protein-Ligand Interactions</i> Author Email(s): sreenivas.sukumar@hpe.com, jacob.balma@hpe.com, avose@aaronvose.net, petersy@muscu.edu, amar@olemiss.edu, ppandey@olemiss.edu, cryates2@olemiss.edu, ikhan@olemiss.edu Contact Person: Sreenivas Sukumar <sreenivas.sukumar@hpe.com>
P285	Shaudi Mahdavi Hosseini, <i>Learning framework with joint task optimization applied to consumer health applications with behavioral nudges</i> Author Email(s): shaudmh@gmail.com Contact Person: Shaudi Mahdavi Hosseini <shaudmh@gmail.com>
P286	Ahmad Qamar, Mohd. Omar, and Muhammad Rashid, <i>An Indexing Scheme for Telerehabilitation Big Data</i> Author Email(s): amq13_com012@student.usm.my, adib@usm.my, mfelahi@uqu.edu.sa Contact Person: Ahmad Qamar <amq13_com012@student.usm.my>
P287	Derrick Ming Yang Lee, Dion Wei Xuan Ang, Grace Mei Ching Pua, Lee Ning Ng, Sharon Purbowo, Eugene Wen Jia Choy, and Kyong Jin Shim, <i>A Social Network Analysis of Jobs and Skills</i> Author Email(s): kjshim@smu.edu.sg Contact Person: Kyong Jin Shim <kjshim@smu.edu.sg>
P288	Xuesi Sim, Ern Rae Chang, Yu Xiang Ong, Jie Ying Yeo, Christine Bai Shuang Yan, Eugene Wen Jia Choy, and Kyong Jin Shim, <i>Digital Social Listening on Conversations About Sexual Harassment</i> Author Email(s): kjshim@smu.edu.sg Contact Person: Kyong Jin Shim <kjshim@smu.edu.sg>
P289	Sebastian Zhi Tao Khoo, Leong Hock Ho, Ee Hong Lee, Danston Kheng Boon Goh, Zehao Zhang, Swee Hong Ng, Haodi Qi, and Kyong Jin Shim, <i>Social Media Analytics: A Case Study of Singapore General Election 2020</i> Author Email(s): kjshim@smu.edu.sg Contact Person: Kyong Jin Shim <kjshim@smu.edu.sg>
P290	Colin Berry, Heather Berry, and Ryan Berry, <i>Mask Mandates and COVID-19 Infection Growth Rates</i> Author Email(s): berryh@gwu.edu, berrycolin17@gmail.com, rberry39@gt.edu Contact Person: Heather Berry <berryh@gwu.edu>