# IEEE BigData 2022 Program Schedule

## Osaka, Japan
Conference time zone: Japanese Standard Time (JST)

The emails with the login information to access the BigData 2022 virtual platform are going to sent out all the registered participants on around Dec 12, 2022. Pls check your spam folder if you haven't receive the email by then. If you need the email login information to access the virtual platform pls email Stephen Woods swoods@computer.org , don't forget to attach the registration receipt if you ask for the login information.   For any questions about the BigData 2022 virtual platform, please email Stephen Woods <swoods@computer.org>

# *Sponsored by*

*Platinum Sponsor:*



**KPMG Consulting Co., Ltd.**

*Gold Sponsor:*



**Ankura**

# HITACHI
## Inspire the Next

**Hitachi**

**NTT DaTa** NTT DATA INTELLILINK Corporation

**NTT Data Intellilink Corporation**

## Think in Data

**Think in Data Initiative, Association Inc**

**Cactus Communications K.K.**

**FRONTEO**

**Google**

**PwC Consulting LLC**



**ZIONEX**

*Bronze Sponsor:*

- Artificial Intelligence Research Center
- Avgidea, Inc.
- Cygames, Inc.
- Iret, Inc.
- Karakuri Inc.
- National Institute of Advanced Industrial Science and Technology
- The Prudential Life Insurance Company, Ltd.
- Trust Architecture Inc.
- VALUENEX Japan Inc.
- Avgidea, Inc.,
- Karakuri Inc.,

# Table of Contents

# IEEE Big Data 2022 Program Schedule

Osaka, Japan
December 17 - December 20, 2022

---

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 Japan Standard time.

---

# Registration: Dec 16  15:00-19:00PM 12th Floor Foyer

| Day 1: Saturday, December 17, 2022 | | | |
|---|---|---|---|
| **Time** | **Location** | **Workshops, Special Sessions, Tutorials, Big Data Cup Challenges** | **Workshop Chair** |
| 8:15-19:00 | 12th Floor Foyer | Registration | |
| 8:00-19:00 | ONLINE | 8th Special Session on Intelligent Data Mining | Uraz Yavanoglu urazyavanoglu@gmail.com; uraz@gazi.edu.tr; |
| 10:30-12:30 | ONLINE | Tutorial 1: High Performance BigData Processing with Dask and Spark using MPI | Dhabaleswar K. (DK) Panda panda@cse.ohio-state.edu Aamir Shafi shafi.16@osu.edu |
| 9:30 – 12:30 | HYBRID Room 1001 | IEEE BigData Cup: Trip Destination Prediction Challenge 2022 – | Yanbo Pang pybdtc@iis.u-tokyo.ac.jp, Yoshihide Sekimoto sekimoto@csis.u-tokyo.ac.jp, Takehiro Kashiyama t.kashiyama@osaka-ue.ac.jp |
| 9:00−12:30 | Room 1002 | Workshop: 7th Workshop on Real-time Stream Analytics, Stream Mining, CER/CEP & Stream Data Management in Big Data | Sabri Skhiri, sabri.skhiri@euranova.eu |
| 9:00−12:30 | Room 801-2 | 3rd International Workshop on User Understanding from Big Data | Claire Ding, qiding.purdue@gmail.com |
| 9:00−12:30 | Room 1003 | The 6th International Workshop on Big Data Analytic for Cybercrime Investigation and Prevention | Andrii Shalaginov, andrii.shalaginov@kristiania.no |

| | | | |
|---|---|---|---|
| 9:00-12:30 | HYBRID Room 1007 | 1st International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2022)<br>2nd International Workshop on Big Data Analytics for Sustainability (BDA4S 2022)<br>Creating Big Data for morals, ethics, and values in natural language in an automatically inferrable format | Stephen McGough, stephen.mcgough@newcastle.ac.uk |
| 8:30-12:30 | ONLINE | Workshop on Big Data Analytic in Healthcare) | Alex Kuo, akuo@uvic.ca |
| 8:30-12:30 | ONLINE | The 6th Workshop on Graph Techniques for Adversarial Activity Analytics (GTA3 2022) | Jiejun Xu, jxu@hrl.com |
| 8:30-12:30 | ONLINE | The 4th International Workshop on Big Data Tools, Methods, and Use Cases for Innovative Scientific Discovery (BTSD) | Sangkeun Lee, lees4@ornl.gov |
| 8:30-12:30 | ONLINE | The 6th IEEE International Workshop on Big Data for Financial News and Data) | Quanzhi Li, quanzhi.li@gmail.com |
| 8:30-12:30 | ONLINE | Second Workshop on Knowledge Graphs and Big Data) | Yuan An, ya45@drexel.edu |
| 8:30-12:30 | ONLINE | 6th Workshop on Human-in-the-loop Methods and Future of Work in BigData) | Atsuyuki Morishima, morishima-office@ml.cc.tsukuba.ac.jp |
| 8:30-12:30 | ONLINE | IEEE BigData 2022 Workshop on Knowledge Discovery and Data Mining in IT Operations (BigData-IT) ) | Yu, Deng, dengy@us.ibm.com, Meng Jiang mjiang2@nd.edu |
| 8:30-12:30 | ONLINE | The 7th IEEE International Workshop on Big Spatial Data (BSD 2022) | Ashwin Shashidharan, ashashidharan@esri.com, FARNOUSH.BANAEI-KASHANI@ucdenver.edu, hendawi@uri.edu, cyzhang@microsoft.com |
| 8:30-12:30 | ONLINE | 1st International Workshop on Responsible AI and Data Ethics (RAIDE 2022) | Ying Xie, yxie2@kennesaw.edu |
| 8:30-12:30 | ONLINE | Creating Big Data for morals, ethics, and values in natural language in an automatically inferrable format | Keisuke Nakamura, keisuke@marusei-sk.com |
| 8:30-12:30 | ONLINE | IEEE Big Data Workshop on Multimodal Big Data (MMBD 2022) | Duoduo Liao <dliao2@gmu.edu> |
| 8:30-12:30 | ONLINE | The 6th Workshop on Benchmarking, Performance Tuning and Optimization for Big Data Applications (BPOD) | Zhiyuan Chen, zhchen@umbc.edu |
| 12:30-14:00 | **Lunch (provided by conference)** | | |
| 14:00-16:00 | ONLINE | Tutorial 2: Explainable AI: Bridging the AI decisions and human trust | Sonali Agarwal sonali@iiita.ac.in<br>Sanjay Kumar Sonbhadra sanjaykumarsonbhadra@soa.ac.in<br>Narinder Singh Punn pse2017002@iiita.ac.in |
| 14:00 – 18:30 | HYBRID Room 1001 | IEEE BigData Cup: Vehicle class and Orientation Detection Challenge 2022 | Ashutosh Kumar ashutosh@iis.u-tokyo.ac.jp, Hiroya Maeda maedahi@iis.u- |

| | | | |
|---|---|---|---|
| | | | tokyo.ac.jp, Hiroshi Omata homata@iis.u-tokyo.ac.jp, Takehiro Kashiyama t.kashiyama@osaka-ue.ac.jp, Yoshihide Sekimoto sekimoto@csis.u-tokyo.ac.jp |
| 14:00 – 18:30 | ONLINE | 5th Special Session on HealthCare Data in IEEE Big Data 2022 | Özgün PINARER opinarer@gsu.edu.tr SULTAN N. TURHAN <sturhan@gsu.edu.tr> |
| 16:00 – 19:00 | ONLINE | 8th Special Session on Information Granulation in Data Science and Scalable Computing | Shusaku Tsumoto tsumoto@med.shimane-u.ac.jp Dominik Slezak slezak@mimuw.edu.pl Tzung-Pei Hong tphong@nuk.edu.tw Leon S. L. Wang slwang@nuk.edu.tw Weiping Ding dwp9988@hotmail.com |
| 14:00-18:30 | HYBRID Room 1007 | 1st International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2022) 2nd International Workshop on Big Data Analytics for Sustainability (BDA4S 2022) Creating Big Data for morals, ethics, and values in natural language in an automatically inferrable format | Stephen McGough, stephen.mcgough@newcastle.ac.uk |
| 14:00-18:30 | Room 1002 | Workshop: Big Data Analytics for Humanitarian Crises | Thilanka Munasinghe, munast@rpi.edu |
| 14:00-18:30 | Room 1003 | The 6th International Workshop on Big Data Analytic for Cybercrime Investigation and Prevention | Andrii Shalaginov, andrii.shalaginov@kristiania.no |
| 14:00-18:30 | ONLINE | 5th Workshop on Big Data for CyberSecurity (BigCyber-2022) | Sudip Mittal, mittal@cse.msstate.edu |
| 14:00-18:30 | ONLINE | The Fifth Workshop on Big Data for Economic and Business Forecasting | Wei Shang, shangwei@amss.ac.cn |
| 14:00-18:30 | ONLINE | The 5th International Workshop on Big Media Dataset Construction, Management and Applications | Cheng Jin, jc@fudan.edu.cn |
| 14:00-18:30 | ONLINE | Creating Big Data for morals, ethics, and values in natural language in an automatically inferrable format | Keisuke Nakamura, keisuke@marusei-sk.com |

| Day 2: Sunday, December 18, 2022 | |
|---|---|
| 8:15-18:00 | **Registration (12th floor foyer)** |
| 8:40-9:15am | **Opening Ceremony** Conference Chair: Shusaku Tsumoto (Izumo) , tsumoto@med.shimane-u.ac.jp; |

| | | | |
|---|---|---|---|
| | PC Chairs : Yukio Ohsawa, ohsawa@sys.t.u-tokyo.ac.jp, Lei Chen, leichen@cse.ust.hk;, Dirk Van den Poel, Dirk.VandenPoel@ugent.be;<br>I&G Chairs: Takuya Takagi, takagi.takuya@fujitsu.com; Yoichi Motomura, 'y.motomura@aist.go.jp';<br>Lingfei Wu,'lwu@email.wm.edu'<br>Steering Committee Chair: Hu,Xiaohua Tony xh29@drexel.edu<br>(Room Event Hall A +12F Conference Room (Broadcast) ) | | |
| 9:15 -10:15 am | **Keynote Speech 1:**<br>Prof. Renee Miller,<br>University Northeastern University,<br>A Vision for Data Alignment and Integration in Data Lakes<br>Chair: Yukio Ohsawa / Dirk Van den Poel<br>ohsawa@sys.t.u-tokyo.ac.jp / Dirk.VandenPoel@ugent.be<br>Dec 18: 9:15 in JST<br>(Room Event Hall A + 12F Conference Hall (Broadcast)) | | |
| 10:15-10:30 am | **Coffee Break** | | |
| 10:00-12:30 | **Poster Setup**<br>**(Room Event Hall BC)** | | |

| Time | Location | Sessions/Workshops | Session chair |
|---|---|---|---|
| 10:30 – 12:30 | Room 1003 | S1: Deep learning for Big Data (1) | Ordonez, Carlos,carlos@central.uh.edu |
| 10:30 – 12:30 | Room 1004-5 | S2: Mining Data from Dynamic Real Society | Yanbo Pang, ybdtc@iis.u-tokyo.ac.jp |
| 10:30 – 12:30 | Room 1009 | S3: Methods for Search and Mining from Big Data (1) | Hiroyuki Kitagawa, kitagawa@cs.tsukuba.ac.jp |
| 10:30 – 12:30 | Room 1002 | S4: Novel Theoretical Models for Big Data (1) | Weimao Ke, weimao.ke@gmail.com |
| 10:30 – 12:30 | Room 1008 | S5: Cloud/Grid/Stream Computing for Big Data | Pranava Mummoju, pranava.mummoju@univie.ac.at |
| 10:30 – 12:30 | Room 801-2 | S6: Predictive analytics on Big Data (1) | U KANG, ukang@snu.ac.kr |
| 10:30 – 12:30 | Conference Hall | I&G S1: Deep Learning 1 | Kiyoshi Izumi izumi@sys.t.u-tokyo.ac.jp |
| 10:30 – 12:30 | HYBRID Room 1001 | Tutorial 3: Spatial Big Data Meets AI: Applications and Perspectives | Zipei Fan fanzipei@iis.u-tokyo.ac.jp Wenlu Wang wenlu.wang@tamucc.edu |
| 10:30 – 12:30 | HYBRID Event Hall A | IEEE Big Data Cup: Privacy-preserving Matching of Encrypted Images | Dominik Slezak slezak@mimuw.edu.pl, Marcin Szczuka m.szczuka@myled.pl, Andrzej Janusz andrzej.janusz@qed.pl |
| 10:30 – 12:30 | HYBRID Room 1007 | IEEE BigData Cup: Crowd sensing-based Road Damage Detection Challenge | Deeksha Arya deeksha.arya1211@gmail.co |

| | | | m, Takehiro Kashiyama t.kashiyama@osaka-ue.ac.jp |
|---|---|---|---|
| 9:30-12:30 | Room1006 | Ninth International Workshop on High Performance Big Graph Data Management, Analysis, and Mining (BigGraphs 2022) | Kames h Madduri, madduri@psu.edu |
| 8:30-12:30 | ONLINE | The 2nd International Workshop on Multi-Modal Medical Data Analysis | Bo Zhang, zbo@bupt.edu.cn |
| 8:30-12:30 | ONLINE | Fifth International Workshop on the Internet of Things Data Analytics (IoTDA) | Eyhab Al-Masri, ealmasriu@gmail.com |
| 8:30-12:30 | HYBRID Room 805 | The 1st International Workshop on Neural Network Models for Brain Connectome Analysis | Carl Yang, j.carlyang@emory.edu |
| 8:30-12:30 | ONLINE | 1st International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2022) 2nd International Workshop on Big Data Analytics for Sustainability (BDA4S 2022) Creating Big Data for morals, ethics, and values in natural language in an automatically inferrable format | Stephen McGough, stephen.mcgough@newcastle.ac.uk |
| 12:30 – 14:00 | **Lunch (provided by conference)** 12:40-14:40 Seminar by KPMG (Room 1003) | | |
| 14:00 – 16:00 | Room 1003 | S7: Deep learning for Big Data (2) | Chong Li ch.l@huawei.com |
| 14:00 – 16:00 | Room 1004-5 | S8: Mining Data from Physical and Cyber Society | Yicheng Tu tuy@usf.edu |
| 14:00 – 16:00 | Room 1009 | S9: Methods for Search and Mining from Big Data (2) | Lamine Diop lamine.diop@epita.fr |
| 14:00 – 16:00 | Room 1002 | S10: Novel Theoretical Models for Big Data (2) | Naoki Watanabe naoki50@keio.jp |
| 14:00 – 16:00 | Room 1008 | S11: Systems to Support Big Data Computing (1) | Yusuke Tanimura yusuke.tanimura@aist.go.jp |
| 14:00 – 16:00 | Room 801-2 | S12: Predictive analytics on Big Data (2) | Takashi Kido kido.takashi@gmail.com |
| 14:00 – 16:00 | Conference Hall | I&G S2 Deep Learning (2) + I&G S3: Graph Analysis (1) | Keiichi Zempo zempo@iit.tsukuba.ac.jp Munehiko Sasajima sasajima@gk.u-hyogo.ac.jp |
| 14:00 – 16:00 | Room 804 | Special Session on Platform for DFFT (Data Free Flow with Trust) | Noboru Koshizuka, noboru@koshizuka-lab.org |
| 14:00 – 16:00 | Room 1006 | Workshop: Applications of Big Data Technology in the Transport Industry | John Easton, j.m.easton@bham.ac.uk |
| 14:00 – 16:00 | ONLINE | 5th Special Session on HealthCare Data in IEEE Big Data 2022 | Özgün PINARER opinarer@gsu.edu.tr SULTAN N. TURHAN <sturhan@gsu.edu.tr> |

| | | | |
|---|---|---|---|
| 14:00 – 16:00 | ONLINE | Tutorial 4: Measurement Informatics | Takashi Washio washio@ar.sanken.osaka-u.ac.jp |
| 14:00 – 16:00 | Hybrid Room 1007 | IEEE BigData Cup: Crowd sensing-based Road Damage Detection Challenge | Deeksha Arya deeksha.arya1211@gmail.com, Takehiro Kashiyama t.kashiyama@osaka-ue.ac.jp |
| 16:00-16:20 | **Coffee Break** | | |
| 16:20 – 18:20 | Room 1003 | S13: Deep learning for Big Data (3) | Rao, Aniruddha Rajendra Aniruddha.Rao@hal.hitachi.com |
| 16:20 – 18:20 | Room 1004-5 | S14: Link and Graph Mining | Yasufumi Takama ytakama@tmu.ac.jp |
| 16:20 – 18:20 | Room 1009 | S15: Complex Big Data Applications (1) | Akinori Abe ave@chiba-u.jp |
| 16:20 – 18:20 | Room 1002 | S16: Computational Models in Big Data | John Paisley jpaisley@columbia.edu |
| 16:20 – 18:20 | Room 1008 | S17: Systems to Support Big Data Computing (2) | Toshiyuki AMAGASA amagasa@cs.tsukuba.ac.jp |
| 16:20 – 18:20 | Room 801-2 | S18: Machine learning algorithms for Big Data (1) | Dirk Van den Poel dirk.vandenpoel@ugent.be |
| 16:20 – 18:20 | Conference Hall | I&G S4: Graph Analysis (2) + I&G S5: Clustering | Takuya Takagi takagi.takuya@fujitsu.com |
| 16:20 – 18:20 | Room 1007 | 8th Special Session on Information Granulation in Data Science and Scalable Computing | Shusaku Tsumoto tsumoto@med.shimane-u.ac.jp Dominik Slezak slezak@mimuw.edu.pl Tzung-Pei Hong tphong@nuk.edu.tw Leon S. L. Wang slwang@nuk.edu.tw Weiping Ding dwp9988@hotmail.com |
| 16:20 – 18:20 | Room 1006 | Workshop: Applications of Big Data Technology in the Transport Industry | John Easton, j.m.easton@bham.ac.uk |
| ONLINE Workshop 14:00-18:30 | ONLINE | 5th Special Session on HealthCare Data in IEEE Big Data 2022 | Özgün PINARER opinarer@gsu.edu.tr SULTAN N. TURHAN <sturhan@gsu.edu.tr> |
| | ONLINE | 8th International Workshop on Methods to Improve Big Data Science Projects (MIDP-2022) | Jeffrey Saltz, jsaltz@syr.edu |
| | ONLINE | Workshop: Scalable Cloud Data Management | Benjamin Wollmer, benjamin.wollmer@uni-hamburg.de |

| | | | |
|---|---|---|---|
| | ONLINE | The 3rd International Workshop on Big Data Reduction (IWBDR-3) | Dingwen Tao, dingwen.tao@wsu.edu |
| 14:00-18:30 | ONLINE | 1st International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2022)<br>2nd International Workshop on Big Data Analytics for Sustainability (BDA4S 2022)<br>Creating Big Data for morals, ethics, and values in natural language in an automatically inferrable format | Stephen McGough, stephen.mcgough@newcastle.ac.uk |
| | ONLINE | 11th Workshop on Scalable Cloud Data Management | Benjamin Wollmer <benjamin.wollmer@uni-hamburg.de> |
| 11:30OPM(Dec 18)– 3:30AM (Dec 19) | ONLINE | Workshop: The Sixth Annual Workshop on Applications of Artificial Intelligence in the Legal Industry<br>https://ankura.zoom.us/webinar/register/WN_WdXc7S_FRfaeUWXSViGKSg | Han Qin, han.qin@ankura.com |

| Day 3: Monday, December 19, 2022 | | | |
|---|---|---|---|
| 8:15-18:00 | Registration (12th floor foyer) | | |
| 9:00 – 10: 00 am | **Keynote Speech 2:**<br>Prof. Hiroaki Kitano,<br>Okinawa Institute of Science and Technology (OIST)<br>COVID-19 AI & Simulation Project: Overview and Lessons Learned<br>Chair: Yukio Ohsawa / Dirk Van den Poel<br>Dec 19: 9:00 in JST<br>(Event Hall A + 12F Conference Hall (Broadcast) ) | | |
| 10: 00-10:30 am | **Coffee Break** | | |
| **Time** | **Location** | **Sessions/Workshops** | **Session chair** |
| 10:30 – 12:30 | Room 1003 | S19: Complex Big Data Applications (2) | Kaira Sekiguchi kaira@sys.t.u-tokyo.ac.jp |
| 10:30 – 12:30 | Room 1009 | S20: Deep learning for Big Data (3) | Takafumi Koshinaka <koshinak@yokohama-cu.ac.jp> |
| 10:30 – 12:30 | Room 1004-5 | S21: Machine learning algorithms for Big Data (2) | Akinori Abe ave@chiba-u.jp |
| 10:30 – 12:30 | Room 1002 | S22: Data Acquisition, Pre-processing, and Integration (1) | Lalana Kagal lkagal@csail.mit.edu |
| 10:30 – 12:30 | Room 801-2 | S23: Big Data in Real Life Applications (1) | FARAH AL.SHANIK falshan@g.clemson.edu |

| 10:30 – 12:30 | Room 1008 | S24: Data Market and Trust | Masatoshi Yoshikawa yoshikawa@i.kyoto-u.ac.jp |
|---|---|---|---|
| 10:30 – 12:30 | Event Hall A | I&G S6: Big Data Systems (1) | Takuya Takagi takagi.takuya@fujitsu.com |
| 10:30 – 12:30 | Conference Room | I&G S7: Legal applications + I&G S8: Covid-19 Analysis | Kurihara Satoshi(IG7) satoshi@keio.jp Bradley Eck (IG8) <bradley.eck@ie.ibm.com> |
| 10:30 – 12:30 | Room 1001 | Special Session on Data-Driven Designation and Implementation of Automated Guided Vehicles | Jerry Chun-Wei Lin jerrylin@ieee.org Rafal Cupek Dariusz Mrozek Gautam Srivastava |
| 10:30 – 12:30 | Room 1006 | Special session on Machine Leaning on Big Data | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it Mst Shapna Akter – Jannatul.shapna99@gmail.com , Renhe Jiang – jiangrh@csis.u-tokyo.ac.jp Emrul Hasan – e1hasan@ryerson.ca Hsing-Kuo Pao – pao@mail.ntust.edu.tw Tom Dooney – tom.dooney@ou.nl , |
| 10:30 – 12:30 | Room 1007 | Special session on Privacy and Security of Big Data (Room 1007) | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it , Rudolf Mayer - rmayer@sba-research.org Muhammad Mazhar Ullah Rathore - rathore.mazhar@unb.ca Nazmus Sakib - nsakib1@kennesaw.edu Jeremy Wu - jeremywu@ucsb.edu Sanguthevar Rajasekaran - sanguthevar.rajasekaran@uconn.edu |
| 10:30-12:30 | ONLINE | Tutorial 5: Learning with Time Series Data: Algorithms, Explainability, and Applications | Jun Huan lukehuan@amazon.com Yuyang Wang yuyawang@amazon.com Youngsuk Park pyoungsu@amazon.com Karthick Gopalswamy kgopalsw@amazon.com Gaurav Gupta |

| | | | gauravaz@amazon.com Panpan Xu xupanpan@amazon.com |
|---|---|---|---|
| 8:30-12:30 | ONLINE | Computational Archival Science: digital records in the age of big data | Mark Hedges, mark.hedges@kcl.ac.uk |
| 8:30-12:30 | ONLINE | 10th International Workshop on Distributed Storage and Blockchain Technologies for Big Data | Hui Li (lih64@pkusz.edu.cn, huilihuge@163.com, Han Wang, wanghan2017@pku.edu.cn |
| 8:30-12:30 | Room 804 | 7th Workshop on Advances in High Dimensional (AdHD) Big Data | Aristidis Vrahatis aris.vrahatis@gmail.com |
| 8:30-12:30 | Room 805 | Second International Workshop on Data science for equality, inclusion and well-being challenges (DS4EIW 2022) | Tania Cerquitelli, tania.cerquitelli@polito.it |
| 8:30-12:30 | ONLINE | 1st International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2022) 2nd International Workshop on Big Data Analytics for Sustainability (BDA4S 2022) Creating Big Data for morals, ethics, and values in natural language in an automatically inferrable format | Stephen McGough, stephen.mcgough@newcastle.ac.uk |
| 12:30 – 14:00 | **Lunch (provided by Conference)** <br> **Poster Display (Event Hall BC)** | | |
| 14:00 – 16:00 | Room 1003 | S25: Complex Big Data Applications (3) | Wenbin Gan wenbingan@nict.go.jp |
| 14:00 – 16:00 | Room 1009 | S26: Deep learning for Big Data (4) | Takafumi Koshinaka <koshinak@yokohama-cu.ac.jp> |
| 14:00 – 16:00 | Conference Room | S27: Machine learning algorithms for Big Data (3) | Takashi Kido of Teikyo University (kido.takashi@gmail.com |
| 14:00 – 16:00 | Room 1002 | S28: Data Acquisition, Pre-processing, and Integration (2) | Hassan Eldeeb ha.eldeeb@gmail.com |
| 14:00 – 16:00 | Room 104-5 | S29: Visualization Analytics for Big Data | Sangbong Yoo usangbong@gmail.com |
| 14:00 – 16:00 | Room 1008 | S30: Trust, resilience, privacy, and security issues | Yang Cao yang@ist.hokudai.ac.jp |
| 14:00 – 16:00 | Event Hall A | (I&G S9: Advertisement and Recommendation + I&G S10: Big Data Analysis (1) | Keisuke Yamazaki k.yamazaki@aist.go.jp |
| 14:00 – 16:00 | Room 1006 | Special session on Machine Leaning on Big Data | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it |
| 14:00−16:00 | Room 1007 | Special session on Privacy and Security of Big Data | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it |
| 14:00−16:00 | ONLINE | Tutorial 6: Utility of Clustering algorithms in Social Media Analysis | Takeaki Uno uno@nii.jp |
| 14:00−16:00 | ONLINE | 5th Special Session on HealthCare Data in IEEE Big Data 2022 | Özgün PINARER opinarer@gsu.edu.tr SULTAN N. TURHAN <sturhan@gsu.edu.tr> |
| 14:00−16:00 | Room 1001 | Special Session on Data-Driven Designation and Implementation of Automated Guided Vehicles | Jerry Chun-Wei Lin jerrylin@ieee.org Rafal Cupek Dariusz Mrozek Gautam Srivastava |

| | | | |
|---|---|---|---|
| 14:00 – 16:00 | ONLINE | IEEE Big Data Cup: GAN-Based Training for Binary Classifier | Hussam Al Hamadi hussam.alhamadi@ku.ac.ae, Ernesto Damiani Ernesto.damiani@ku.ac.ae, Martina Morcos martina.morcos@ku.ac.ae,, Sivaprasad Nandyala siva@ssrc.tii.ae, Brian McGillion brian@ssrc.tii.ae |
| 14:00-18:00 | Room 801-2 | (Room 801-2) Workshop: New Strands of Usage of Big Data in Medical Systems, Market and Institutional Design, and Economic Theory | Naoki Watanabe, naoki50@keio.jp |
| 14:00-18:00 | Room 805 | (Room 805) Workshop on Cyber Threat Intelligence and Hunting (CyberHunt 2022) | Vasileios Mavroeidis, vasileim@ifi.uio.no |
| 14:00-18:00 | ONLINE | Computational Archival Science: digital records in the age of big data | Mark Hedges, mark.hedges@kcl.ac.uk |
| 14:00-18:00 | ONLINE | 10th International Workshop on Distributed Storage and Blockchain Technologies for Big Data | Han Wang, wanghan2017@pku.edu.cn |
| 14:00-18:00 | ONLINE | Big Data Engineering and Analytics in Cyber-Physical Systems" (BigEACPS'22) | Akbar Namin, akbar.namin@ttu.edu |
| 14:00-18:00 | ONLINE | Big Food, Nutrition and Environment Data Management and Analysis | Barbara Koroušić Seljak, barbara.korousic@ijs.si |
| 16: 00-16:30 | **Coffee Break** | | |
| 16:30-18:50 | ONLINE | Special session on Machine Leaning on Big Data  - | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it, Mst Shapna Akter – Jannatul.shapna99@gmail.com, Renhe Jiang – jiangrh@csis.u-tokyo.ac.jp Emrul Hasan – e1hasan@ryerson.ca Hsing-Kuo Pao – pao@mail.ntust.edu.tw Tom Dooney – tom.dooney@ou.nl |
| 16:30-18:50 | ONLINE | Special session on Privacy and Security of Big Data -- | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it, Rudolf Mayer - rmayer@sba-research.org Muhammad Mazhar Ullah Rathore - rathore.mazhar@unb.ca Nazmus Sakib - nsakib1@kennesaw.edu Jeremy Wu - jeremywu@ucsb.edu Sanguthevar Rajasekaran - sanguthevar.rajasekaran@uconn.edu |
| 16:30-18:50 | ONLINE | Computational Archival Science: digital records in the age of big data | Mark Hedges, mark.hedges@kcl.ac.uk |
| 16:30-18:50 | ONLINE | 10th International Workshop on Distributed Storage and Blockchain Technologies for Big Data | Han Wang, wanghan2017@pku.edu.cn |

| | |
|---|---|
| 16:30 – 17:30 | **Keynote Speech 3:**<br>Prof. Meihui Zhang<br>Beijing Institute of Technology, China<br>Big Data Management and Analytics for Supporting Smart Healthcare Applications<br>Chair: Shusaku Tsumoto<br>tsumoto@med.shimane-u.ac.jp<br>Dec 19: 16:30 in JST (AM 7:30 by GST)<br>(Event Hall A+12F Conference Hall (Broadcast)) |
| 18:00-21:00 | **Banquet**<br>Award Ceremony<br>Social Program<br>(RIHGA ROYAL HOTEL:<br>next to Conference Venue) |

| Day 4: Tuesday, December 20, 2022 | | | |
|---|---|---|---|
| 8:30-12:30 | **Registration (12th floor foyer)** | | |
| 9:00 - 10: 00 | **Keynote Speech 4:**<br>Prof. Tamer Özsu<br>A SYSTEMATIC VIEW OF DATA SCIENCE<br>University of Waterloo (2018)<br>Chair: Dirk van den Poel<br>Dirk.VandenPoel@ugent.be<br>Dec 20: 9:00 in JST<br>(Room Event Hall A + 12F Conference Hall (Broadcast)) | | |
| 10:00 – 10:30 | **Coffee Break** | | |
| **Time** | **Location** | **Sessions/Workshops** | **Session chair** |
| 10:30 – 12:30 | Room 1001 | S31: Trust, resilience, privacy, and security issues | Min-Te Sun<br>msun@csie.ncu.edu.tw |
| 10:30 – 12:30 | Room 1002 | S32: Social Web Search and Mining | Lamine Diop<br>lamine.diop@epita.fr |
| 10:30 – 12:30 | Event Hall A | S33: Big Data in Real Life Applications (2) | John Easton<br>j.m.easton@bham.ac.uk |
| 10:30 – 12:30 | Room 1009 | S34: Large-scale Recommendation Systems and Social Media Systems | Daichi AMAGATA<br>amagata.daichi@ist.osaka-u.ac.jp |
| 10:30 – 12:30 | Room 1004-5 | S35: Big Data Analytics in Government, Public Sector and Society in General | Roten, Daniel<br>d1roten@sdsc.edu |
| 10:30 – 12:30 | Room 1003 | S36: Deep learning for Big Data (5) | Jaewoo Lee<br>JAEWOO.LEE@uga.edu |

| Time | Room | Session | Contact |
|---|---|---|---|
| 10:30 – 12:30 | Room 801-2 | S37: Learning Essential Features in Big Data | Rao, Aniruddha Rajendra Aniruddha.Rao@hal.hitachi.com |
| 10:30 – 12:30 | Conference Hall | I&G S11: Big Data Analysis (2) | Swati Agarwal <swatia@goa.bits-pilani.ac.in> |
| 10:30 – 12:30 | Room 1008 | I&G S12: Big Data Systems (2) + I&G S13: Big Data Applications (1) + I&G S14: Big Data Analysis (3) | Hong-Linh Truong <linh.truong@aalto.fi> Keisuke Yamazaki k.yamazaki@aist.go.jp |
| 10:30 – 12:30 | Room 1006 | Special session on Machine Leaning on Big Data | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it |
| 10:30 – 12:30 | Room 1007 | Special session on Privacy and Security of Big Data | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it |
| 8:30 – 12:30 | Room 805 | Workshop: Digital twins for accelerated discovery of climate and sustainability solutions | Levente Klein, kleinl@us.ibm.com |
| 8:30 – 12:30 | Room 804 | Workshop: The Sixth Annual Workshop on Applications of Artificial Intelligence in the Legal Industry | Han Qin, han.qin@ankura.com |
| 8:30-12:30 | ONLINE | 1st International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2022) 2nd International Workshop on Big Data Analytics for Sustainability (BDA4S 2022) Creating Big Data for morals, ethics, and values in natural language in an automatically inferrable format | Stephen McGough, stephen.mcgough@newcastle.ac.uk |
| 12:30 – 14:00 | Lunch (provided by the Conference) Poster Display (Event Hall BC) | | |
| 14:00 – 16;00 | Room 1001 | S38: Data Ecosystem and Its Evaluation (1) | Kawsar Haghshenas kawsar.haghshenas@iaas.uni-stuttgart.de |
| 14:00 – 16;00 | Room 1002 | S39: Social Web Search and Mining | Hiroki Sakaji sakaji@sys.t.u-tokyo.ac.jp |
| 14:00 – 16;00 | Event Hall A | S40: Big Data in Real Life Applications (3) | Yukio Ohsawa ohsawa@sys.t.u-tokyo.ac.jp |
| 14:00 – 16;00 | Room 1009 | S41: Large-scale Recommendation Systems and Social Media Systems | Katsutoshi Yada yada@kansai-u.ac.jp |
| 14:00 – 16;00 | Room 1004-5 | S42: Big Data Analytics for Value Creation in the Real Life | Bessho Masahiro bessho@iniad.org |
| 14:00 – 16;00 | Room 1003 | S43: Deep learning for Big Data (6) | Poonam Goyal drpoonamgoyal@gmail.com |
| 14:00 – 16;00 | Conference Hall | I&G S15: Big Data Applications (2) | Tsukasa Ishigaki <isgk@tohoku.ac.jp> |
| 14:00 – 16;00 | Hybrid Room 1008 | (Room 1008) Tutorial 7: Learning from Non-IID Data: Centralized vs. Federated Learning | Jun Wu junwu3@illinois.edu Wenxuan Bao wbao4@illinois.edu Jingrui He jingrui@illinois.edu |
| 16:00 - 16: 20 | **Coffee Break** | | |
| 16:20-18:00 | Room 1003 | S44: Link and Graph Mining | Shen Wang <wssddtc@gmail.com> |

| | | | |
|---|---|---|---|
| 16:20-18:00 | Room 1001 | S45: Computational Models and Information Quality | Ordonez, Carlos <carlos@Central.UH.EDU> |
| 16:20-18:00 | Conference Hall | I&G S16: Big Data Analysis (4) | Takuya Takagi takagi.takuya@fujitsu.com |
| 16:20-18:00 | ONLINE | Special session on Machine Leaning on Big Data  - | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it, Mst Shapna Akter – Jannatul.shapna99@gmail.com, Renhe Jiang – jiangrh@csis.u-tokyo.ac.jp Emrul Hasan – e1hasan@ryerson.ca Hsing-Kuo Pao – pao@mail.ntust.edu.tw Tom Dooney – tom.dooney@ou.nl |
| 16:20-18:00 | ONLINE | Special session on Privacy and Security of Big Data | Alfredo Cuzzocrea alfredo.cuzzocrea@unical.it, Rudolf Mayer - rmayer@sba-research.org Muhammad Mazhar Ullah Rathore - rathore.mazhar@unb.ca Nazmus Sakib - nsakib1@kennesaw.edu Jeremy Wu - jeremywu@ucsb.edu Sanguthevar Rajasekaran - sanguthevar.rajasekaran@uconn.edu |
| 18:00-18:30 | **Closing (12F Conference Hall)** | | |

# Main Conference Paper Sessions

| S1: Deep learning for Big Data (1) | |
|---|---|
| BigD254 "FEZ: a Flexible and Efficient Zoom-in Method for Ultra-large Image Classification" (Regular) | Cheng-Han Lu and Shih-Hao Hung |
| BigD286 "Improving Representation Learning for Session-based Recommendation" (Regular) | Tianwen Chen and Raymond Chi-Wing Wong |
| BigD246 "FedLesScan: Mitigating Stragglers in Serverless Federated Learning" (Short) | Mohamed Elzohairy, Mohak Chadha, Anshul Jindal, Andreas Grafberger, Jianfeng Gu, Michael Gerndt, and Osama Abboud |
| BigD250 "Open Set Domain Adaptation by Extreme Value Theory" (Short) | Yiming Xu and Diego Klabjan |
| BigD308 "VWC-BERT:Scaling Vulnerability–Weakness–Exploit Mapping on Modern AI Accelerators" (Short) | Siddhartha Shankar Das, Mahantesh Halappanavar, Antonino Tumeo, Edoardo Serra, Alex Pothen, and Ehab Al-Shaer |
| BigD309 "Sequential Anomaly Detection with Local and Global Explanations" (Short) | He Cheng, Depeng Xu, and Shuhan Yuan |

| S2: Mining Data from Dynamic Real Society | |
|---|---|
| BigD614 "Clustering Human Mobility with Multiple Spaces" (Regular) | Haoji Hu, Haowen Lin, and Yao-Yi Chiang |
| BigD563 "Proximal Cost-sensitive Sparse Group Online Learning" (Regular) | Zhong Chen, Huixin Zhan, Victor Sheng, Andrea Edwards, and Kun Zhang |
| BigD413 "On Inferring User Socioeconomic Status with Mobility Records" (Regular) | Zheng Wang, Mingrui Liu, Cheng Long, Qianru Zhang, Jiangneng Li, and Chunyan Miao |
| BigD831 "Spatial Attention Based Grid Representation Learning For Predicting Origin–Destination Flow" (Regular) | Mingfei Cai, Yanbo Pang, and Yoshihide Sekimoto |
| BigD598 "Generation of Synthetic Urban Vehicle Trajectories" (Short) | Chrysovalantis Anastasiou, Seon Ho Kim, and Cyrus Shahabi |

| S3: Methos for Search and Mining from Big Data (1) | |
|---|---|
| BigD366 "LASSO Logic Engine: harnessing the logic parsing capabilities of the LASSO algorithm for longitudinal feature learning" (Regular) | Jason Orender, Mohammad Zubair, and Jiangwen Sun |
| BigD633 "VIMTS: Variational-based Imputation for Multi-modal Time Series" (Regular) | Xiaohu Zhao, Kebin Jia, Benjamin Letcher, Jennifer Fair, Yiqun Xie, and Xiaowei Jia |
| BigD262 "Multi-Query Optimization Revisited: A Full-Query Algebraic Method" (Regular) | Mehrad Eslami, Yicheng Tu, Zichen Xu, and Hadi Charkhgard |
| BigD472 "A New IoT Storage System Based on Raw NVM" (Short) | Tao Cai, Yueming Ma, Dejiao Niu, Pengfei Gao, Tianle Lei, and Jianfei Dai |
| BigD650 "An FPGA-based Accelerator for Regular Path Queries over Edge-labeled Graphs" (Short) | Kento Miura, Ryohei Kobayashi, Toshiyuki Amagasa, Hiroyuki Kitagawa, Norihisa Fujita, and Taisuke Boku |
| BigD651 "Verification of Sparsity in the Attention Mechanism of Transformer for Anomaly Detection in Multivariate Time Series" (Short) | Chihiro Maru, Boris Brandherm, and Ichiro Kobayashi |

| S4: Novel Theoretical Models for Big Data (1) | |
|---|---|
| BigD310 "Trie-based Output Space Itemset Sampling" (Regular) | Lamine Diop, Cheikh Talibouya Diop, Arnaud Giacometti, and Arnaud Soulet |
| BigD550 "Alternatives to Classic BM25-IDF based on a New Information-theoretical Framework" (Regular) | Weimao Ke |

| | |
|---|---|
| BigD240 "Adaptive and Dynamic Multi-Resolution Hashing for Pairwise Summations" (Short) | Lianke Qin, Aravind Reddy, Zhao Song, Zhaozhuo Xu, and Danyang Zhuo |
| BigD442 "A Novel GPU-Accelerated Algorithm to Discover Periodic-Frequent Patterns in Temporal Databases" (Short) | Uday Kiran Rage, Tarun Sreepada, Yutaka Watanobe, and Kazuo Goda |
| BigD647 "FVec2vec: A Fast Nonlinear Dimensionality Reduction Approach for General Data" | Xiaoli Ren, Kefeng Deng, Kaijun Ren, Junqiang Song, Xiaoyong Li, and Qing Xu |
| **S5: Cloud/Grid/Stream Computing for Big Data** | |
| BigD743 "Ruya: Memory-Aware Iterative Optimization of Cluster Configurations for Big Data Processing" (Regular) | Jonathan Will, Lauritz Thamsen, Jonathan Bader, Dominik Scheinert, and Odej Kao |
| BigD411 "Multi-Cloud Query Optimisation with Accurate and Efficient Quoting" (Short) | Damien T. Wojtowicz, Shaoyi Yin, Jorge Martinez-Gil, Franck Morvan, and Abdelkader Hameurlain |
| BigD450 "PROFET: Profiling-based CNN Training Latency Prophet for GPU Cloud Instances" (Short) | Sungjae Lee, Yoonseo Hur, Subin Park, and Kyungyong Lee |
| BigD497 "Latency vs Cost : Data Ingestion options at Twitter Scale" (Short) | Zhenzhao Wang, Santosh Marella, Mak Inada, Pablo Rodriguez Defino, Abhishek Jagannath, and Lohit VijayaRenu |
| BigD710 "Enhancing k-Means Algorithm with Tensor Processing Unit" (Short) | Pranava Mummoju, Anna Wolff, Martin Perdacher, Claudia Plant, and Christian B¨ohm |
| BigD816 "Optimal Rate Control for Latency-constrained High Throughput Big Data Applications" (Short) | Ziren Xiao, Aaron Harwood, and Maria Rodriguez |
| BigD849 "Perona: Robust Infrastructure Fingerprinting for Resource-Efficient Big Data Analytics" (Short) | Dominik Scheinert, Soeren Becker, Jonathan Bader, Lauritz Thamsen, Jonathan Will, and Odej Kao |
| **S6: Predictive analytics on Big Data (1)** | |
| BigD372 "Tab-HGNN: Learning Column Representation with Heterogeneous Graph Neural Network for Web Table Interpretation" (Regular) | Xingyu Su and Yanyan Shen |
| BigD448 "A Unified Model for Bid Landscape Forecasting in the Mixed Auction Types of Real-Time Bidding" (Regular) | Seonguk Seo, Jihye Ha, Jieun Shin, Sunah Kim, and Taeho Hwang |
| BigD471 "Accurate PARAFAC2 Decomposition for Temporal Irregular Tensors with Missing Values" (Regular) | Jun-Gi Jang, Jeongyoung Lee, Jiwon Park, and U Kang |
| BigD656 "Spatio-Temporal Based Architecture Topology Search for Multivariate Time Series Prediction" (Short) | Xinqi Lyu, Yibo Chen, Jiamin Chen, Xiangyue Liu, and Jianliang Gao |
| BigD498 "A Semi-Supervised Learning Approach for Abnormal Event Prediction on Large Network Operation Time-Series Data" (Regular) | Yijun Lin and Yao-Yi Chiang |
| **S7: Deep learning for Big Data (2)** | |
| BigD295 "Learning Dynamic Multimodal Implicit and Explicit Networks for Multiple Financial Tasks" (Regular) | Gary Ang and Ee-Peng Lim |
| BigD318 "Decomposed Transformer with Frequency Attention for Multivariate Time Series Anomaly Detection" (Regular) | Shuxin Qin, Jing Zhu, Dan Wang, Liang Ou, Hongxin Gui, and Gaofeng Tao |
| BigD339 "WaveNets: Wavelet Channel Attention Networks" (Regular) | Hadi Salman, Caleb Parks, Shi Hong, and Justin Zhan |
| BigD412 "Enhancing Graph Convolutional Networks by Topology Sampling" (Short) | Quentin Petit, Chong Li, Serge Petiton, Kelun Chai, and Nahid Emad |
| BigD425 "SHCN: Self-supervised General Hypergraph Clustering Network" (Short) | Mengran Li, Xinlin Piao, Yong Zhang, Yongli Hu, and Baocai Yin |
| **S8: Mining Data from Physical and Cyber Society** | |
| BigD566 "iFiG: Individually Fair Multi-view Graph Clustering" (Regular) | Yian Wang, Jian Kang, Yinglong Xia, Jiebo Luo, and Hanghang Tong |

| | |
|---|---|
| BigD649 "ABM: Attention-based Message Passing Network for Knowledge Graph Completion" (Regular) | Weikai Xu, Lihui Liu, and Hanghang Tong |
| BigD283 "Towards Robust Graph Neural Networks via Adversarial Contrastive Learning" (Regular) | Shen Wang, Zhengzhang Chen, Jingchao Ni, Haifeng Chen, and Philip S. Yu |
| BigD815 "FedGSL: Federated Graph Structure Learning for Local Subgraph Augmentation" (Short) | Geng Zhao, Yi Huang, Chihsu Tsai, and Yuxia Sun |
| BigD452 "QuoGNN: Quotient Graph Neural Network for Urban Flow Forecasting" (Short) | Xiaochuan Gou, Peng Han, and Xiangliang Zhang |
| **S9: Methods for Search and Mining from Big Data (2)** | |
| BigD617 "Predicting Deduplication Performance: An Analytical Model and Empirical Evaluation" (Regular) | Owen Randall and Paul Lu |
| BigD237 "Scalable and Accurate Density-Peaks Clustering on Fully Dynamic Data" (Regular) | Daichi Amagata |
| BigD215 "A Novel Estimate Fitness Method by Measuring A large number of Sub-event Logs Conformance" (Short) | Mimi Wang, Xudong He, and Peihai Zhao |
| BigD771 "From Data Warehouse to Lakehouse: A Comparative Review" (Short) | Ahmed Harby and Farhana zulkernine |
| BigD575 "Griddle: Effective Query Support over Voluminous Gridded Spatial Datasets" (Short) | Pierce Smith, Sangmi Pallickara, and Shrideep Pallickara |
| **S10:  Novel Theoretical Models for Big Data (2)** | |
| BigD571 "On High-Latency Bowtie Data Streaming" (Regular) | Gabriel Stella and Dmitri Loguinov |
| BigD688 "Detection of Unobserved Common Cause in Discrete Data Based on the MDL Principle" (Regular) | Masatoshi Kobayashi, Kohei Miyaguchi, and Shin Matsushima |
| BigD755 "Improving Robustness of Vision Transformers via Data-Augmented Virtual Adversarial Training" (Short) | Weili Shi and Sheng Li |
| **S11: Systems to Support Big Data Computing (1)** | |
| BigD445 "ByteStore: Hybrid Layouts for Main-Memory Column Stores" (Regular) | Pengfei Zhang, ZiQiang Feng, Eric Lo, and Hailin Qin |
| BigD505 "Scalable Parallel Machine Learning Computing a Summarization Matrix with SQL Queries" (Regular) | Carlos Ordonez |
| BigD422 "Workload Prediction for Adaptive Approximate Query Processing" (Short) | Hamid Shahrivari, Odysseas Papapetrou, and George Fletcher |
| BigD454 "A novel self-adaptive method for improving patient monitoring with composite early-warning scores" (Short) | Antonio Iyda Paganelli, Pedro Elkind Velmovitsky, Adriano Branco, Markus Endler, Plinio Pelegrini Morita, Paulo Alencar, and Donald Cowan |
| **S12: Predictive analytics on Big Data (2)** | |
| BigD552 "Profiling-free Configuration Adaptation and Latency-Aware Resource Scheduling for Video Analytics" (Regular) | Tian Zhou, Fubao Wu, and Lixin Gao |
| BigD758 "Significant Digits: Using Large-Scale Blockchain Data to Predict Fraudulent Addresses" (Regular) | Jared Gridley and Oshani Seneviratne |
| BigD845 "Bundle entropy as an optimized measure of consumers' systematic product choice combinations in mass transactional data" (Regular) | Roberto Mansilla, Gavin Smith, Andrew Smith, and James Goulding |
| BigD787 "Fidelity and diversity metrics for validating hierarchical synthetic data: Application to residential energy demand" (Short) | Swapna Thorve, Anil Vullikanti, Henning Mortveit, Samarth Swarup, and Madhav Marathe |
| **S13: Deep learning for Big Data (3)** | |

| | |
|---|---|
| BigD415 "A Hybrid Model Based on NeuralProphet and Long Short-Term Memory for Time Series Forecasting" (Regular) | Zhaofeng Yu, Kun Niu, Xiao Chen, Zhijun Guo, and Dongsheng Li |
| BigD392 "Local Temporal Abstraction in Time Series Classification" (Short) | Nevo Itzhak, Shahar Tal, Hadas Cohen, Osher Daniel, Roze Kopylov, and Robert Moskovitch |
| **S14: Link and Graph Mining (8-8)** | |
| BigD577 "Skew-Symmetric Adjacency Matrices for Clustering Directed Graphs" (Regular) | Koby Hayashi, Sinan Aksoy, and Haesun Park |
| BigD628 "Self-supervised Hypergraph Representation Learning" (Regular) | Boxin Du, Changhe Yuan, Robert Barton, Tal Neiman, and Hanghang Tong |
| BigD687 "Dynamic Graph Summarization: Optimal and Scalable" (Regular) | Mahdi Hajiabadi, Venkatesh Srinivasan, and Alex Thomo |
| BigD453 "Fast Binary Network Hashing via Graph Clustering" (Short) | Yasuhiro Fujiwara, Masahiro Nakano, Atsutoshi Kumagai, Yasutoshi Ida, Akisato Kimura, and Naonori Ueda |
| BigD330 "DELATOR: Money Laundering Detection via Multi-Task Learning on Large Transaction Graphs" (Short) | Henrique S. Assumpção, Fabrício Souza, Leandro Lacerda Campos, Vinícius T. de Castro Pires, Paulo M. Laurentys de Almeida, and Fabricio Murai |
| BigD464 "Network Distance based on Laplacian Flows on Graphs" (Short) | Dianbin Bao, Kisung You, and Lizhen Lin |
| **S15: Complex Big Data Applications (1)** | |
| BigD233 "Revenue Maximization in Waterfall Auctions with Dynamic Programming" (Regular) | Dan Halbersberg, Matan Halevi, and Moshe Salhov |
| BigD243 "Conditional Graph Regression for Complex Chemical Systems with Heterogeneous Substructures" (Regular) | Gyoung S. Na |
| BigD328 "Progressive prediction of hospitalisation and patient disposition in the emergency department" (Regular) | Laura Uhl, Vincent AUGUSTO, Vincent Lemaire, Youenn Alexandre, Fanny Jardinaud, Paolo Bercelli, and Saber Aloui |
| BigD245 "POSGen: Personalized Opening Sentence Generation for Online Insurance Sales" (Short) | Yu Li, Yi Zhang, Weijia Wu, Zimu Zhou, and Qiang Li |
| BigD311 "ISIDE: Proactively Assist University Students at Risk of Dropout" (Short) | Enrico Bassetti, Andrea Conti, Emanuele Panizzi, and Gabriele Tolomei |
| BigD320 "Continuous Health Monitoring of Machinery using Online Clustering on Unlabeled Data Streams" (Short) | Minh-Huong Le-Nguyen, Fabien Turgis, Pierre-Emmanuel Fayemi, and Albert Bifet |
| **S16: Computational Models in Big Data** | |
| BigD390 "Aggregation Delayed Federated Learning" (Regular) | Ye Xue, Diego Klabjan, and Yuan Luo |
| BigD590 "Efficient first-order predictor-corrector multiple objective optimization for fair misinformation detection" (Regular) | Eric Enouen, Katja Mathesius, Sean Wang, Arielle Carr, and Sihong Xie |
| BigD695 "Probabilistic Orthogonal Matching Pursuit" (Regular) | Ghazal Fazelnia and John Paisley |
| BigD726 "Bayesian Nonparametric Model Averaging Using Scalable Gaussian Process Representations" (Regular) | John Paisley, Sebastian Rowland, Jeremiah Liu, Brent Coull, and Marianthi Kioumourtzoglou |
| BigD434 "Scalable Tensors for Big Data Analytics" (Short) | Leonidas Fegaras, Tanvir Ahmed Khan, Md Hasanuzzaman Noor, and Tanzima Sultana |
| **S17: Systems to Support Big Data Computing (2)** | |
| BigD591 "Data Transformation Acceleration using Deterministic Finite-State Transducers" (Regular) | Marziyeh Nourian, Tri Nguyen, Andrew A. Chien, and Michela Becchi |

| BigD299 "Adaptive Aggregation For Federated Learning" (Short) | K. R. Jayaram, Vinod Muthusamy, Gegi Thomas, Ashish Verma, and Mark Purcell |
|---|---|

| **S18: Machine learning algorithms for Big Data (1)** | |
|---|---|
| BigD256 "Exploring Multi-Time Context Vector and Randomness for Stock Movement Prediction" (Regular) | Kanghyeon Seo and Jihoon Yang |
| BigD260 "LLU-Swin:Low-Light Image Enhancement with U-shaped Swin Transformer" (Regular) | Lijiang Shao, Wenjin Wang, Shiguo Huang, and Xiaolin Li |
| BigD302 "AuxPOS: Improving Grammatical Correctness with Big Data Based Text Summarization" (Regular) | Rui Zhang, Jianjun Yu, and Yuanchun Zhou |
| BigD365 "Spatial Clustering Method of Historical AIS Data for Maritime Traffic Routes Extraction" (Regular) | Lubna Eljabu, Mohammad Etemad, and Stan Matwin |
| BigD297 "A Convergence Theory for Federated Average: Beyond Smoothness" (Short) | Xiaoxiao Li, Zhao Song, Runzhou Tao, and Guangyi Zhang |
| BigD377 "Adaptive Knowledge Transfer on Evolving Domains" (Short) | Jun Wu, Hanghang Tong, Elizabeth Ainsworth, and Jingrui He |

| **S19: Complex Big Data Applications (2)** | |
|---|---|
| BigD444 "Encouraging Crowd Avoidance Behavior using Dynamic Pricing Framework Towards Preventing the Spread of COVID-19" (Regular) | Keiichi Ochiai, Hiroyuki Sato, Hiroshi Kawakami, Takahiro Ide, Toru Otaki, Akira Yamada, Tatsuya Yano, Hiroki Okawa, Takuya Shirai, and Yutaka Arakawa |
| BigD579 "Multi-Behavior Recommendation with Hyperbolic Geometry" (Regular) | Di You, Thanh Tran, and Kyumin Lee |
| BigD604 "ESEP: Data-Driven Emergency and Safe Evacuation Driving Path Planning During Natural Catastrophes" (Regular) | Shohaib Mahmud, Haiying Shen, Ying Natasha Zhang Foutz, and Joshua Anton |
| BigD327 "Leveraging Subgraph Structure for Exploration and Analysis of Bitcoin Address" (Short) | Yuexin Xiang, Tiantian Li, and Yuquan Li |
| BigD337 "Secure Multi-party kNN Search in Large-scale Spatial Data Federation" (Short) | Yuanyuan Zhang, Yexuan Shi, Nan Zhou, Yi Xu, and Ke Xu |
| BigD396 "Teachers in Social Media: A Gender-aware Behavior Analysis" (Short) | Hamid Karimi, Kaitlin Torphy Knake, and Kenneth A. Frank |

| **S20: Deep learning for Big Data (3)** | |
|---|---|
| BigD421 "Heterogeneity-Aware Adaptive Federated Learning Scheduling" (Regular) | Jingoo Han, Ahmad Faraz Khan, Syed Zawad, Ali Anwar, Nathalie Baracaldo Angel, Yi Zhou, Feng Yan, and Ali R. Butt |
| BigD527 "EFFGAN: Ensembles of fine-tuned federated GANs" (Regular) | Ebba Ekblom, Edvin Listo Zec, and Olof Mogren |
| Fucxxy kBigD536 "Randomize Adversarial Defense in a Light Way" (Regular) | Ji-Young Park, Lin Liu, Jixue Liu, and Jiuyong Li |
| BigD443 "Dynamic Cell Structure via Recursive-Recurrent Neural Networks" (Short) | Xin Qian, Matthew Kennedy, and Diego Klabjan |
| BigD484 "Road Rutting Detection using Deep Learning on Images" (Short) | Poonam Kumari Saha, Deeksha Arya, Ashutosh Kumar, Hiroya Maeda, and Yoshihide Sekimoto |
| BigD516 "Asymmetric Self-Supervised Graph Neural Networks" (Short) | Zhuo Tan, Bin Liu, and Guosheng Yin |

| **S21: Machine learning algorithms for Big Data (2)** | |
|---|---|
| BigD490 "Efficient and Scalable Implicit Graph Neural Networks with Virtual Equilibrium" (Regular) | Qi Chen, Yifei Wang, Yisen Wang, Jianlong Chang, Qi Tian, Jiansheng Yang, and Zhouchen Lin |
| BigD587 "GQFormer: A Multi-Quantile Generative Transformer for Time Series Forecasting" (Regular) | Shayan Jawed and Lars Schmidt-Thieme |
| BigD630 "Robust and Resource-efficient Machine Learning Aided Viewport Prediction in Virtual Reality" (Regular) | Yuang Jiang, Konstantinos Poularakis, Diego Kiedanski, Sastry Kompella, and Leandros Tassiulas |

| | |
|---|---|
| BigD564 "Fuzzy Law: Towards Creating a Novel Explainable Technology-Assisted Review System for e-Discovery" (Short) | Charles Courchaine and Ricky Sethi |
| BigD569 "Robust Learning with Noisy Label Detection and Counterfactual Correction" (Short) | Wenting Qi and Charalampos Chelmis |
| BigD574 "Label Denoising and Counterfactual Explanation with A Plug and Play Framework" (Short) | Wenting Qi and Charalampos Chelmis |

| S22: Data Acquisition, Pre-processing, and Integration | |
|---|---|
| BigD348 "An approach to constructing a graph data repository for course recommendation based on IT career goals in the context of big data" (Regular) | Thu Nguyen, Ngoc Vu, and Binh Ly |
| BigD462 "Overcoming Challenges of Synthetic Data Generation" (Regular) | Kevin Fang, Vaikkunth Mugunthan, Vayd Ramkumar, and Lalana Kagal |
| BigD689 "Two-Stage Sampling: A Framework for Imbalanced Classification With Overlapped Classes" (Regular) | Neni Alya Firdausanti, Tirana Noor Fatyanosa, Mahendra Data, Israel Mendonca, and Masayoshi Aritsugi |
| BigD489 "Data Virtual Machines: Simplifying Data Sharing, Exploration & Querying in Big Data Environments" (Short) | Damianos Chatziantoniou, Verena Kantere, Nikos Antoniou, and Angeliki Gantzia |
| BigD819 "Efficient Privacy Preserving Record Linkage at Scale using Apache Spark" (Short) | Alexandros Karakasidis and Georgia Koloniari |
| BigD332 "Clustering-based Unsupervised Generative Relation Extraction" (Short) | Chenhan Yuan, Ryan Rossi, Andrew Katz, and Hoda Eldardiry |

| S23: Big Data in Real Life Applications (1) | |
|---|---|
| BigD500 "Semantic-based Attention model for Hospital Readmission Prediction" (Short) | Sea Jung Im, Yue Xu, and Jason Watson |
| BigD502 "Graph Structure Neural Differential Equations on Spatio-temporal Prediction" (Short) | Hangtao He and Kejiang Ye |
| BigD511 "Hypergraph Learning based Recommender System for Anomaly Detection, Control and Optimization" (Short) | Sakhinana Sagar Srinivas, Rajat Kumar Sarkar, and Venkataramana Runkana |
| BigD523 "Auto-Regressive Self-Attention Models for Diagnosis Prediction on Electronic Health Records" (Short) | Pascal Wullschleger, Simone Lionetti, Donnacha Daly, Francesca Volpe, and Grégoire Caro |
| BigD531 " Towards Automatic Cetacean Photo-Identification: A Framework for Fine-Grain, Few-Shot Learning in Marine Ecology" (Short) | Cameron Trotter, Nick Wright, A. Stephen McGough, Matt Sharpe, Barbara Cheney, Mònica Arso Civil, Reny Tyson Moore, Jason Allen, and Per Berggren |
| BigD619 "Evaluating Distribution System Reliability with Hyperstructures Graph Convolutional Nets" (Short) | Yuzhou Chen, Tian Jiang, Miguel Heleno, Alexandre Moreira, and Yulia R. Gel |

| S24: Data Market and Trust | |
|---|---|
| BigD723 "A Task-Driven Privacy-Preserving Data-Sharing Framework For Industrial Internet" (Regular) | Parshin Shojaee, Yingyan Zeng, Muntasir Wahed, Avi Seth, Ran Jin, and Ismini Lourentzou |
| BigD829 "FL-Market: Trading Private Models in Federated Learning" (Regular) | Shuyuan Zheng, Yang Cao, Masatoshi Yoshikawa, Huizhong Li, and Qiang Yan |
| BigD388 "Robust Text CAPTCHAs Using Adversarial Examples" (Regular) | Rulin Shao |
| BigD537 "SMARDY: Zero-Trust FAIR Marketplace for Research Data" (Short) | Ion Dorinel Filip, Cosmin Ionite, Alba Gonzalez-Cebrian, Mihaela Balanescu, Ciprian Dobre, Adriana E. Chis, Dave Feenan, Adrian-Alexandru Buga, Ioan-Mihai Constantin, George Suciu, George Iordache, and Horacio Gonzalez-Velez |
| BigD273 "Efficient Soft-Error Detection for Low-precision Deep Learning Recommendation Models" (Short) | Sihuan Li and Zizhong Chen |

| | |
|---|---|
| BigD441 "RAF: Recursive Adversarial Attacks on Face Recognition Using Extremely Limited Queries" (Short) | Keshav Kasichainula, Hadi Mansourifar, and Weidong Shi |

| **S25: Complex Big Data Applications (3)** ||
|---|---|
| BigD611 "Accurate Stock Movement Prediction with Self-supervised Learning from Sparse Noisy Tweets" (Regular) | Yejun Soun, Jaemin Yoo, Minyong Cho, Jihyeong Jeon, and U Kang |
| BigD735 "APIP: Attention-based Protein Representation Learning for Protein-Ligand Interface Prediction" (Regular) | Tian Xia, Bo Hui, and Wei-Shinn Ku |
| BigD397 "Penalty Prediction Based on Modulated Hierarchical Attention Coupled with Legal Attribute Recognition" (Short) | Jingzhe Liu and Peng Wu |
| BigD401 "An Open Case-based Reasoning Framework for Personalized On-board Driving Assistance in Risk Scenarios" (Short) | Wenbin Gan, Minh-Son Dao, and Koji Zettsu |
| BigD460 "Automatic Eating Behavior Detection from Wrist Motion Sensor Using Bayesian, Gradient Boosting, and Topological Persistence Methods" (Short) | Yu-Min Chung, Amir Nikooienejad, and Bo Zhang |

| **S26: Deep learning for Big Data (4)** ||
|---|---|
| BigD540 "RGB Color Model Aware Computational Color Naming and Its Application to Data Augmentation" (Regular) | Zipei Yan, Linchuan Xu, Atsushi Suzuki, Jing Wang, Jiannong Cao, and Jun Huang |
| BigD601 "Time Series Anomaly Detection using Prediction-Reconstruction Mixture Errors" (Regular) | Lawrence Wong, Dongyu Liu, Laure Berti-Equille, and Kalyan Veeramachaneni |
| BigD624 "DFFNet: Dynamic Feature Fusion Network for Weakly Supervised Object Detection in Remote Sensing Images" (Short) | Mengmeng Zhu, Shouhong Wan, Peiquan Jin, and Peiyi Zhang |
| BigD625 "Exploring the Target Distribution for Surrogate-Based Black-Box Attacks" (Short) | Raha Moraffah, Paras Sheth, and Huan Liu |
| BigD742 "Continuous Prompt Tuning Based Textual Entailment Model for E-commerce Entity Typing" (Short) | Yibo Wang, Congying Xia, Guan Wang, and Philip Yu |
| BigD585 "Forecasting Emerging Pandemics with Transfer Learning and Location-aware News Analysis" (Regular, video-presentation) | Jing Chen, German Creamer, and Yue Ning |

| **S27: Machine learning algorithms for Big Data (3)** ||
|---|---|
| BigD699 "AnyStreamKM: Anytime k-medoids Clustering for Streaming Data" (Regular) | Jagat Sesh Challa, Dhruv Rawat, Navneet Goyal, and Poonam Goyal |
| BigD785 "Semi-supervised Embedding for Scalable and Accurate Time Series Clustering" (Regular) | Andrew Hill, Russell Bowler, Katerina Kechris, and Farnoush Banaei-Kashani |
| BigD813 "Multiple Tiered Treatments Optimization with Causal Inference on Response Distribution" (Regular) | Tao Huang, Qingyang Li, and Zhiwei Qin |
| BigD613 "Fast Support Vector Machine Using Singular Value Decomposition" (Short) | Sabuj Laskar and Muhammad Abdullah Adnan |
| BigD776 "Critic-over-Actor-Critic Modeling: Finding Optimal Strategy in ICU Environments" (Short) | Riazat Ryan and Ming Shao |
| BigD814 "Uncertainty Aware Trader-Company Method: Interpretable Stock Price Prediction Capturing Uncertainty" (Short) | Yugo Fujimoto, Kei Nakagawa, Kentaro Imajo, and Kentaro Minami |

| **S28: Data Acquisition, Pre-processing, and Integration** ||
|---|---|
| BigD728 "BigFeat: Scalable and Interpretable Automated Feature Engineering Framework" (Regular) | Hassan Eldeeb, Shota Amashukeli, and Radwa Elshawi |
| BigD580 "Operator Placement for Spatio-temporal Tasks" (Regular) | Daniel Glake, Mareike Schmidt, Felix Kiehn, Fabian Panse, Ulfia Lenfers, Thomas Clemen, and Norbert Ritter |

| | |
|---|---|
| BigD717 "GenSyn: A Multi-stage Framework for Generating Synthetic Microdata using Macro Data Sources" (Short) | Angeela Acharya, Siddhartha Sikdar, Sanmay Das, and Huzefa Rangwala |
| BigD721 "Generating Interpretable Features for Context-Aware Document Clustering: A Cybersecurity Case Study" (Short) | Luis Felipe Gutiérrez and Akbar Namin |
| BigD565 "Visualizing and Querying Large-scale Structured Datasets by Learning Multi-layered 3D Meta-Profiles" (Short) | Michael Gubanov and Sophie PAvia |
| BigD487 "Resource Efficient Profiling of Spatial Variability in Performance of Regression Models" (Short) | Menuka Warushavithana, Saptashwa Mitra, and Shrideep Pallickara |
| BigD533 "Proactive Query Expansion for Streaming Data Using External Sources" (Short) | Farah Alshanik, Amy Apon, Yuheng Du, Alexander Herzog, and Ilya Safro |

| S29: Visual Analytics for Big Data | |
|---|---|
| BigD275 "DCPViz: A Visual Analytics Approach for Downscaled Climate Projections" (Regular) | Abdullah-Al-Raihan Nayeem, Huikyo Lee, Dongyun Han, Mohammad Elshambakey, William J. Tolone, Todd Dobbs, Daniel Crichton, and Isaac Cho |
| BigD646 "Shapelet-based Temporal Association Rule Mining for Multivariate Time Series Classification" (Regular) | Omar Bahri, Soukaina Filali Boubrahimi, and Shah Muhammad Hamdi |
| BigD833 "Visual Analytics System of Comprehensive Data Quality Improvement for Machine Learning using Data- and Process-driven Strategies" (Short) | Hyein Hong, Sangbong Yoo, Yejin Jin, Chanyoung Yoon, Soobin Yim, Seokhwan Choi, and Yun Jang |
| BigD486 "Characterizing Spontaneous Ideation Contest on Social Media: Case Study on the Name Change of Facebook to Meta" (Short) | Kunihiro Miyazaki, Takayuki Uchiba, Haewoon Kwak, and Jisun An |
| BigD554 "Learning on Health Fairness and Environmental Justice via Interactive Visualization" (Short) | Abdullah-Al-Raihan Nayeem, Ignacio Segovia-Dominguez, Huikyo Lee, Dongyun Han, Yuzhou Chen, Zhiwei Zhen, Yulia Gel, and Isaac Cho |

| S30: Trust, resilience, privacy, and security issues | |
|---|---|
| BigD420 "Fraud Detection via Contrastive Positive Unlabeled Learning" (Regular) | Vinay M.S., Shuhan Yuan, and Xintao Wu |
| BigD463 "MulBot: Unsupervised Bot Detection Based on Multivariate Time Series" (Regular) | Lorenzo Mannocci, Stefano Cresci, Anna Monreale, Athina Vakali, and Maurizio Tesconi |
| BigD709 "Defending Evasion Attacks via Adversarially Adaptive Training" (Regular) | Minh-Hao Van, Wei Du, Xintao Wu, Feng Chen, and Aidong Lu |
| BigD685 "Heterogeneous Randomized Response for Differential Privacy in Graph Neural Networks" (Short) | Khang Tran, Phung Lai, Hai Phan, Issa Khalil, Yao Ma, Abdallah Khreishah, My Thai, and Xintao Wu |
| BigD712 "Trustworthy Distributed Deep Neural Network Training in an Edge Devices Network" (Short) | Sudipta Saha Shubha and Haiying Shen |
| BigD559 "Asymmetric Differential Privacy" (Short) | Shun Takagi, Fumiyuki Kato, Yang Cao, and Masatoshi Yoshikawa |

| S31: Trust, resilience, privacy, and security issues | |
|---|---|
| BigD249 "User-Entity Differential Privacy in Learning Natural Language Models" (Regular) | Phung Lai, NhatHai Phan, Tong Sun, Rajiv Jain, Franck Dernoncourt, Jiuxiang Gu, and Nikolaos Barmpalios |
| BigD756 "Location-based Alert System Using Searchable Encryption with Hilbert Curve Encoding" (Regular) | Po-wei Harn, Sai Deepthi Yeddula, Libo Sun, Min-Te Sun, and Wei-Shinn Ku |

| S32: Social Web Search and Mining | |
|---|---|
| BigD493 "Proactive Prioritization of App Issues via Contrastive Learning" (Regular) | Moghis Fereidouni, Adib Mosharrof, Umar Farooq, and A.B. Siddique |
| BigD603 "Abusive Language Detection with Graph based Multi-task Learning" (Regular) | Chunyun Zhang, Xi Zhang, Quan Wang, Jiayi Liang, Ge Zhang, Sanchuan Guo, Wenyu Zang, and Yongdong Zhang |

| | |
|---|---|
| BigD353 "Emotion-Cause Relationship Between Clauses Prediction: a Novel Method Based on BERT for Emotion-Cause Pair Extraction" (Short) | Rui Fan, Tingting He, Yufan Wang, Jie Hong, and Xinhui Tu |
| BigD766 "HyperTwitter: A Hypergraph-based Approach to Identify Influential Twitter Users and Tweets" (Short) | Lulwah Alkulaib, Abdulaziz Alhamadani, Shailik Sarkar, and Chang-Tien Lu |
| BigD394 "Understanding COVID-19 Vaccine Campaign on Facebook with Minimal Supervision" (Regular , Video presentation) | Tunazzina Islam and Dan Goldwasser |

| **S33: Big Data in Real Life Applications (2)** | |
|---|---|
| BigD737 "Transition-Aware Multi-Activity Knowledge Tracing" (Regular) | Siqian Zhao, Chunpai Wang, and Shaghayegh Sahebi |
| BigD751 "Shape-based Evaluation of Epidemic Forecasts" (Regular) | Ajitesh Srivastava, Satwant Singh, and Fiona Lee |
| BigD789 "Enhancing COVID-19 Ensemble Forecasting Model Performance Using Auxiliary Data Sources" (Regular) | Aniruddha Adiga, Gursharn Kaur, Benjamin Hurt, Lijing Wang, Przemyslaw Porebski, Srinivasan Venkatramanan, Bryan Lewis, and Madhav Marathe |
| BigD627 "The impact of product descriptions on customers' trust in sellers and purchase intention" (Short) | Yi Sun and Yukio Ohsawa |
| BigD644 "DL-BERT: a time-aware double-level BERT-style model with pre-training for disease prediction" (Short) | Xianlai Chen, Jiamiao Lin, and Ying An |
| BigD669 "Vehicle re-identification and trajectory reconstruction using multiple moving cameras in the CARLA driving simulator" (Short) | Ashutosh Kumar, Takehiro Kashiyama, Hiroya Maeda, Fan Zhang, Hiroshi Omata, and Yoshihide Sekimoto |

| **S34: Large-scale Recommendation Systems and Social Media Systems** | |
|---|---|
| BigD458 "Learning Diversity Attributes in Multi-Session Recommendations" (Regular) | Nassim Bouarour, Idir Benouaret, and Sihem Amer-Yahia |
| BigD474 "RI-GCN: Review-aware Interactive Graph Convolutional Network for Review-based Item Recommendation" (Regular) | Yijin Cai, Yilei Wang, Weijin Wang, and Wenting Chen |
| BigD407 "SandPiper: A Cost-Efficient Adaptive Framework for Online Recommender Systems" (Short) | prashanth thinakaran, Jashwant Raj Gunasekaran, Mahmut Taylan Kandemir, and Chita R Das |
| BigD467 "Debiasing Graph Transfer Learning via Item Semantic Clustering for Cross-Domain Recommendations" (Short) | Zhi Li, Daichi Amagata, Yihong Zhang, Takahiro Hara, Shuichiro Haruta, Kei Yonekawa, and Mori Kurokawa |
| BigD521 "Question-Attentive Review-Level Recommendation Explanation" (Short) | Trung-Hoang Le and Hady W. Lauw |

| **S35: Big Data Analytics in Government, Public Sector and Society in General** | |
|---|---|
| BigD232 "MRA-DGCN: Multi-Range Attention-Based Dynamic Graph Convolutional Network for Traffic Prediction" (Regular) | Renyi Chen and HUAXIONG YAO |
| BigD281 "Detecting languages in streetscapes using deep convolutional neural networks" (Regular) | You Xuan Thung, Tom Benson, and Nikita Klimenko |
| BigD350 "How to Protect the Public Opinion Against New Types of Bots?" (Regular) | Jan Ludwig Reubold, Stephan Cornelius Escher, Christian Wressnegger, and Thorsten Strufe |
| BigD491 "Confli-T5: An AutoPrompt Pipeline for Conflict Related Text Augmentation" (Short) | Erick Skorupa Parolin, Yibo Hu, Latifur Khan, Patrick T. Brandt, Javier Osorio, and Vito D'Orazio |
| BigD588 "New Technologies, Old Propaganda Strategies: How Russian Disinformation about Biolabs in Ukraine was spread via Conservative and Anti-Vaccination Discourse on Twitter" (Short) | Iuliia Alieva and Kathleen M. Carley |
| BigD659 "Cross-Lingual Text Classification of Transliterated Hindi and Malayalam" (Short) | Jitin Krishnan, Antonios Anastasopoulos, Hemant Purohit, and Huzefa Rangwala |

| S36: Deep learning for Big Data (5) | |
|---|---|
| BigD635 "Physics-based multiple time-series univariate forecasting" (Regular) | Sumanta Mukherjee, Arindam Jati, Kanthi Sarpatwar, and Roman Vaculin |
| BigD653 "Role-Oriented Dynamic Network Embedding" (Regular) | Ting Pan, Wenjun Wang, Minglai Shao, Yueheng Sun, and Pengfei Jiao |
| BigD668 "PromptRGD: Prompt Learning with Relation-aware Gradient Denoising for Low-resource Relation Extraction" (Regular) | Zihang Xu, Luoyi Fu, and Xinbing Wang |
| BigD767 "Cyberbully Detection Using BERT with Augmented Texts" (Short) | Xiaoyu Guo, Usman Anjum, and Justin Zhan |
| BigD773 "MBAG: A Scalable Mini-Block Adaptive Gradient Method for Deep Neural Networks" (Short) | Jaewoo Lee |
| BigD791 "Predictive Masking for Semi-Supervised Graph Contrastive Learning" (Short) | Yufei Jin and Xingquan Zhu |

| S37: Learning Essential Features in Big Data | |
|---|---|
| BigD496 "SCM-VAE: Learning Identifiable Causal Representations via Structural Knowledge" (Regular) | Aneesh Komanduri, Yongkai Wu, Wen Huang, Feng Chen, and Xintao Wu |
| BigD570 "bigg2vec: Fast and Memory-Efficient Representation Learning for Billion-Scale Graphs on a Single Machine" (Regular) | Mangesh Bendre, Mahashweta Das, Vineeth Rakesh, Fei Wang, and Hao Yang |
| BigD507 "A Functional approach for Two Way Dimension Reduction in Time Series" (Regular) | Aniruddha Rajendra Rao, Haiyan Wang, and Chetan Gupta |
| BigD344 "Dynamic network of commodity futures market and systemic risk contribution of key commodities" (Regular) | Ke Huang, Jifeng Sun, Zuominyang Zhang, and Ying Ye |
| BigD817 "Residual-Quantile Adjustment for Adaptive Training of Physics-informed Neural Network" (Regular) | Jiayue Han, Zhiqiang Cai, Zhiyou Wu, and Xiang Zhou |

| S38: Data Ecosystem and Its Evaluation | |
|---|---|
| BigD419 "Fair Collective Classification in Networked Data" (Regular) | Karuna Bhaila, Yongkai Wu, and Xintao Wu |
| BigD461 "Robust Personalized Federated Learning under Demographic Fairness Heterogeneity" (Regular) | Alycia Carey, Wei Du, and Xintao Wu |
| BigD499 "InfoFair: Information-Theoretic Intersectional Fairness" (Regular) | Jian Kang, Tiankai Xie, Xintao Wu, Ross Maciejewski, and Hanghang Tong |
| BigD582 "Fair Regression under Sample Selection Bias" (Regular) | Wei Du, Xintao Wu, and Hanghang Tong |
| BigD716 "CO2 Emission Aware Scheduling for Deep Neural Network Training Workloads" (Short) | Kawsar Haghshenas, Brian Setz, and Marco Aiello |
| BigD440 "Boosting Demographic Fairness of Face Attribute Classifiers via Latent Adversarial Representations" (Short) | Huimin Zeng, Zhenrui Yue, Lanyu Shang, Yang Zhang, and Dong Wang |
| BigD805 "SG-CF: Shapelet-Guided Counterfactual Explanation for Time Series Classification" (Short) | Peiyu Li, Soukaina Boubrahimi, and Shah Hamdi |

| S39: Social Web Search and Mining | |
|---|---|
| BigD634 "Nothing Stands Alone: Leveraging News Relations through a Hypergraph for Fake News Detection" (Regular) | Ujun Jeong, Kaize Ding, Lu Cheng, Ruocheng Guo, Kai Shu, and Huan Liu |
| BigD739 "Spatial Density-based User Identity Linkage across Social Networks" (Regular) | Hui Xue, Bo Sun, Weixuan Mao, Jianjun Lin, Yijing Zhang, Xiaoyu Liu, Xiaodu Yang, and Ziqian Chen |
| BigD802 "MUSCAT: Multilingual Rumor Detection in Social Media Conversations" (Regular) | Md Rabiul Awal, Minh Dang Nguyen, Roy Ka-Wei Lee, and Kenny Tsu Wei Choo |

| S40: Big Data in Real Life Applications (3) | |
|---|---|
| BigD713 "Building the Next-Generation Radio Networks Planning Infrastructure Using Datacubes" (Short) | Peter Baumann, Johannes Boensmann, and Bang Pham Huu |
| BigD746 "Mobile Application Behaviour Anomaly Detection based on API Calls" (Short) | Fernando Kaway Carvalho Ota, Farouk Damoun, Sofiane Lagraa, Enriqueta Patricia Becerra Sanchez, Christophe Atten, Jean Hilger, and Radu State |
| BigD772 "MatPhase: Material phase prediction for Li-ion battery Reconstruction using Hierarchical Curriculum Learning" (Short) | Anika Tabassum, Nikhil Muralidhar, Ramakrishnan Kannan, and Srikanth Allu |
| BigD801 "A Unified Framework to Assess Market Implications of Institutional Investments" (Short) | Taruna Seth and Vipin Chaudhary |
| BigD841 "AE-smnsMLC: Multi-Label Classification with Semantic Matching and Negative Label Sampling for Product Attribute Value Extraction" (Short) | Zhongfen Deng, Wei-Te Chen, Lei Chen, and Philip Yu |
| BigD842 "Deep Learning for Destination Choice Modeling: A Fundamental Approach for National Level People Flow Reconstruction" (Short) | Yanbo Pang and Yoshihide Sekimoto |

| S41: Large-scale Recommendation Systems and Social Media Systems | |
|---|---|
| BigD514 "TimeKit: A Time-series Forecasting-based Upgrade Kit for Collaborative Filtering" (Regular) | Seoyoung Hong, Minju Jo, Seungji Kook, Jaeeun Jung, Hyowon Wi, and Noseong Park |
| BigD652 "Mitigating Frequency Bias in Next-Basket Recommendation via Deconfounders" (Regular) | Xiaohan Li, Zheng Li, Luyi Ma, Kaushiki Nag, Stephen Guo, Philip Yu, and Kannan Achan |
| BigD683 "Sequential Recommendation with Auxiliary Item Relationships via Multi-Relational Transformer" (Regular) | Ziwei Fan, Zhiwei Liu, Chen Wang, Peijie Huang, Peng Hao, and Philip Yu |
| BigD777 "Time-aware Hyperbolic Graph Attention Network for Session-based Recommendation" (Regular) | Xiaohan Li, Yuqing Liu, Zheng Liu, and Philip Yu |
| BigD676 "Neighborhood Preserving Embedding via Capsules for Collaborative Filtering" (Short) | Xuan Li and Li Zhang |

| S42: Big Data Analytics for Value Creation in the Real Life | |
|---|---|
| BigD362 "Detecting Neighborhood Gentrification at Scale via Street-level Visual Data" (Regular) | Tianyuan Huang, Timothy Dai, Zhecheng Wang, Hesu Yoon, Hao Sheng, Andrew Ng, Ram Rajagopal, and Jackelyn Hwang |
| BigD707 "Machine Learning for Improved Post-fire Debris Flow Likelihoods Prediction" (Regular) | Daniel Roten, Jessica Block, Daniel Crawl, Jenny Lee, and Ilkay Altintas |
| BigD479 "Store Congestion Forecast under the Pandemic using Point of Sales Statistics" (Regular) | Masahiro Bessho and Ken Sakamura |
| BigD547 "Cough Classification with Deep Derived Features using Audio Spectrogram Transformer" (Regular) | Julio Valdes, Karim Habashy, Pengcheng Xi, Madison Cohen-McFarlane, Bruce Wallace, Rafik Goubran, and Frank Knoefel |
| BigD214 "MeWP: Meta-Learning based Water-Level Prediction" (Short) | Justin Mao, Oliver Yun, Conner Kim, Haipeng Chang, and Xuehan Sun |

| S43: Deep learning for Big Data (6) | |
|---|---|
| BigD711 "Deep Learning For Time Series Classification Using New Hand-Crafted Convolution Filters" (Regular) | Ali Ismail-Fawaz, Maxime Devanne, Jonathan Weber, and Germain Forestier |
| BigD731 "AdvCat: Domain-Agnostic Robustness Assessment for Cybersecurity-Critical Applications with Categorical Inputs" (Regular) | Helene Orsini, Hongyan Bao, Yujun Zhou, Xiangrui Xu, Yufei Han, Longyang Yi, Wei Wang, Xin Gao, and Xiangliang Zhang |
| BigD752 "IGRP: Iterative Gradient Rank Pruning for Finding Graph Lottery Ticket" (Regular) | Po-wei Harn, Sai Deepthi Yeddula, Bo Hui, Jie Zhang, Libo Sun, Min-Te Sun, and Wei-Shinn Ku |
| BigD811 "Communication-Efficient TeraByte-Scale Model Training Framework for Online Advertising" (Short) | Weijie Zhao, Xuewu Jiao, Mingqing Hu, Xiaoyun Li, Xiangyu Zhang, and Ping Li |

| BigD822 "A Generalized Multimodal Deep Learning Model for Early Crop Yield Prediction" (Short) | Arshveer Kaur, Poonam Goyal, Kartik Sharma, Lakshay Sharma, and Navneet Goyal |
| --- | --- |
| BigD855 "Dynamic Graph Attention Recurrent Network for Traffic Prediction" (Short) | JIA GUO, HUAXIONG YAO, JUNTAO YANG, ZUOQUAN XIE, RENYI CHEN, and MENGLING HU |
| **S44: Link and Graph Mining** | |
| BigD760 "MetaKGRec: Collaborative Meta-Knowledge-Graph Enhanced Recommender System" (Regular) | Liangwei Yang, Shen Wang, Jibing Gong, Shaojie Zheng, Shuying Du, Zhiwei Liu, and Philip S. Yu |
| BigD774 "Temporal Graph Representation Learning via Maximal Cliques" (Regular) | Shima Khoshraftar, Aijun An, and Nastaran Babanejad |
| BigD475 "Temporal Adaptive Aggregation Network for Dynamic Graph Learning" (Short) | Man Wu and Xingquan Zhu |
| BigD542 "Graggle: A Graph-based Approach to Document Clustering" (Short) | Isaiah King and H. Howie Huang |
| BigD596 "Dynamic Graph Node Classification via Time Augmentation" (Short) | Jiarui Sun, Mengting Gu, Chin-Chia Michael Yeh, Yujie Fan, Girish Chowdhary, and Wei Zhang |
| **S45: Computational Models and Information Quality** | |
| BigD798 "Persistence Homology of Proximity Hyper-Graphs for Higher Dimensional Big Data" (Regular) | Rohit Singh and Philip Wilsey |
| BigD803 "Evaluating Captioning Models using Markov Logic Networks" (Short) | Monika Shah, Somdeb Sarkhel, and Deepak Venugopal |
| BigD824 "Appley: Approximate Shapley Value for Model Explainability in Linear Time" (Short) | Md Shafiul Alam and Ying Xie |
| BigD847 "Informative Initialization and Kernel Selection Improves t-SNE for Biological Sequences" (Short) | Prakash Chourasia, Sarwan Ali, and Murray Patterson |

# Industry and Government Program Sessions

| I&G S1:  Deep Learning (1) | |
| --- | --- |
| N242 "Enhanced Video BERT for Fast Video Advertisement Retrieval"(Regular) | Yi Yang, Tan Yu, Jie Liu, Zhipeng Jin, Xuewu Jiao, Yi Li, Shuanglong Li, and Ping Li |
| N292 "Tree-based Text-Vision BERT for Video Search in Baidu Video Advertising"(Regular) | Tan Yu, Jie Liu, Yi Yang, Yi Li, Hongliang Fei, and Ping Li |
| N216 "Minimal Cost Device Calibration in Spectral Analysis via Meta Learning: Towards Efficient Deployment of Deep Neural Networks in Industry"(Regular) | Xiang Xie, Muen Jin, Anqi Chu, and Wilhelm Stork |
| N275 "Collaborative Filtering Guided Deep Reinforcement Learning for Sequential Recommendations"(Short) | Vahid Azizi, Saayan Mitra, and Xiang Chen |
| N230 "Bilateral Trade Flow Prediction by Gravity Graph Auto-encoder"(Short) | Naoto Minakawa, Kiyoshi Izumi, and Hiroki Sakaji |
| N234 "Enabling Real-Time Low-Cost Spectral Analysis on Edge Devices with Deep Neural Networks: a Robust Hybrid Approach"(Short) | Xiang Xie, Tiantian Chen, and Wilhelm Stork |
| N218 "QUBO-inspired Molecular Fingerprint for Chemical Property Prediction"(Poster) | Koichiro Yawata, Yoshihiro Osakabe, Takuya Okuyama, and Akinori Asahara |

| I&G S2: Deep Learning (2) | |
|---|---|
| N240 "ScaDLES: Scalable Deep Learning over Streaming data at the Edge"(Regular) | Sahil Tyagi and Martin Swany |
| N244 "Protein-Protein Interactions (PPIs) Extraction from Biomedical Literature using Attention-based Relational Context Information"(Regular) | Gilchan Park, Sean Mc Corkle, Carlos Soto, Ian Blaby, and Shinjae Yoo |
| N289 "Semantic Segmentation Approaches in Crop Row Detection"(Short) | Rashed Doha, Mohammad Al Hasan, and Sohel Anwar |
| N269 "A Geomagnetic Sensor Dataset for Traffic Flow Prediction"(Poster) | Huanchen WANG, Quanjun Chen, Zheng DONG, Xuan Song, Hao Tian, Donglong YANG, and Manxia LIU |
| I&G S3: Graph analysis (1) | |
| N287 "Sub-Sequence Graph Representation Learning on High Variability Data for Dynamic Risk Prediction in Critical Care"(Regular) | Ankur Teredesai, Sijin Huang, Tucker Stewart, Juhua Hu, Armaan Thakker, Katie Stern, and Grant O'Keefe |
| N283 "StarBridge: a topological multidimensional subgraph analysis to detect fraudulent nodes and rings in telecom network"(Poster) | Pedro Fidalgo, Rui Lopes, and Christos Faloutsos |
| N280 "Comparative Reasoning for Knowledge Graph Fact Checking"(Poster) | Lihui Liu, Houxiang Ji, Jiejun Xu, and Hanghang Tong |
| I&G S4: Graph analysis (2) | |
| N233 "SETN: Stock Embedding Enhanced with Textual and Network Information"(Short) | Takehiro Takayanagi, Hiroki Sakaji, and Kiyoshi Izumi |
| N261 "Stream2Graph: Dynamic Knowledge Graph for Online Learning Applied in Large-scale Network"(Short) | Mariam BARRY, Albert BIFET, Raja CHIKY, Saad EL JAOUHARI, Jacob MONTIEL, Aissa EL OUAFI, Dhiaddedine YOUSFI, Eric GUERIZEC, and Aurel NOBIAL |
| N243 "Graph Representation Learning based Vulnerable Target Identification in Ransomware Attacks"(Short) | Wenqi Wei, Mu Qiao, Eric Butler, and Divyesh Jadav |
| I&G S5: Clustering | |
| N260 "In-DBMS K-means Clustering for Binary Streams"(Regular) | Carlos Garcia-Alvarado and Carlos Ordonez |
| N226 "A Hybrid Score to Optimize Clustering Hyperparameters for Online Search Term Data"(Short) | Ariana Martino and Allison Rossetto |
| N232 "Improving Paper Classification Using Forecasting"(Short) | Jason Zhang, Krerkkiat Chusap, Wei Zhang, and Chang Liu |
| N206 "A Scalable, Geospatial Segmentation Method for Visiting Trend Analysis"(Poster) | Sumit Jain, Kazune Furudo, Yuiko Kumagai, and Satoshi Kamatani |
| I&G S6: Big Data Systems (1) | |
| N204 "Digital Data Platform for Connected Clinical Trials"(Short) | Hui Zhang, Ju Ji, Guangchen Ruan, Neel Patel, Regan Giesting, Leah Miller, Yi Lin Yang, and Jian Yang |
| N231 "Learning skills adjacency representations for optimized reskilling recommendations"(Short) | Saksham Gandhi, Raj Nagesh, and Subhro Das |
| N246 "The Data Recovery Service in NoSQL"(Short) | Hung-Chang Hsiao |
| N251 "A monitoring framework for deployed machine learning models with supply chain applications"(Short) | Bradley Eck, Duygu Kabakci-Zorlu, Yan Chen, France Savard, and Xiaowei Bao |
| N288 "Distributed Incremental Machine Learning for Big Time Series Data"(Short) | Dhaval Salwala, Seshu Tirupathi, Brian Quanz, Wesley Gifford, Stuart Siegel, Vijay Ekambaram, and Arindam Jati |

| N278 "Regression based Optimizer Statistics generation based on q-error"(Short) | Parag Paul, Wenlu Wang, Haiwuan Chen, and Wei-Shinn Ku |
| --- | --- |
| N236 "Recovery-Robust Inquiry to External Services from Stream Processing Application"(Poster) | Yuki Nishiguchi, Hiroaki Aoki, Takafumi Onishi, Julius Michaelis, and Yasuhiko Kanemasa |
| N268 "Smart System for Multi-Cloud Pathways"(Poster) | Indervir Singh Banipal and Shubhi Asthana |
| N208 "DSServe - Data Science using Serverless"(Poster) | Dhaval Patel, Shuxin Li, and Jayant Kalagnanam |

## I&G S7: Legal applications

| N255 "Named Entity Recognition in Long Documents: An End-to-end Case Study in the Legal Domain"(Regular) | Hossein Keshavarz, Zografoula Vagena, Pigi Kouki, Ilias Fountalis, Mehdi Mabrouki, Aziz Belaweid, and Nikolaos Vasiloglou |
| --- | --- |
| N257 "Explainable Text Classification Techniques in Legal Document Review: Locating Rationales without Using Human Annotated Training Text Snippets"(Regular) | Christian Mahoney, Peter Gronvall, Nathaniel Huber-Fliflet, and Jianping Zhang |
| N256 "Integration of Rule-Based Reasoning and Transfer Learning in Legal Document Review"(Short) | Robert Keeling, Ava Guo, Peter Gronvall, Nathaniel Huber-Fliflet, and Jianping Zhang |
| N282 "DedupeGov: A Platform for Deduplication of Large Volumes of Data From Natural and Legal Entities in the Governmental Field"(Short) | Marcos Carvalho, Vitor Mangaravite, Lucas Ponce, Luis Cantelli, Bruno Campoi, Gabriel Nunes, Bruno Paiva, Alberto Laender, and Marcos Gonçalves |

## I&G S8: Covid-19 Analysis

| N252 "Large-Scale Knowledge Synthesis and Complex Information Retrieval from Biomedical Documents"(Short) | Shreya Saxena, Raj Sangani, Siva Prasad, Shubham Kumar, Mihir Athale, Rohan Awhad, and Vishal Vaddina |
| --- | --- |
| N217 "Media Coverage and Public Perception of Distance Learning During the COVID-19 Pandemic: A Topic Modeling Approach Based on BERTopic"(Poster) | Gloria Hristova and Nikolay Netov |
| N219 "The relationship between Twitter sentiment and mobility during the COVID-19 pandemic"(Poster) | Yoshiyuki Suimon, Hiroto Tanabe, and Kiyoshi Izumi |

## I&G S9: Advertisement and Recommendation

| N285 "Boost CTR Prediction for New Advertisements via Modeling Visual Content"(Regular) | Tan Yu, Zhipeng Jin, Jie Liu, Yi Yang, Hongliang Fei, and Ping Li |
| --- | --- |
| N290 "From Billion to Trillion : Feature Fusion Network for Personalized Online Advertising Systems"(Regular) | Weijie Zhao, Peng Yang, Dong Li, Xing Shen, Lin Liu, and Ping Li |
| N279 "Conversion-Based Dynamic-Creative-Optimization in Native Advertising"(Short) | Yohay Kaplan, Yair Koren, Alex Shtoff, Tomer Shadi, and Oren Somekh |
| N253 "Learning from Disagreement for Event Detection"(Short) | Liang Wang, Junpeng Wang, Yan Zheng, Shubham Jain, Chin-Chia Michael Yeh, Zhongfang Zhuang, Javid Ebrahimi, and Wei Zhang |
| N209 "A Hybrid Approach to Similar Stock Recommender with Taxonomy and Behavior Data"(Poster) | Yizhi Mao, Miao Chen, Bill Dorn, and Yan He |

## I&G S10: Big Data Analysis (1)

| N210 "Better Skill-based Job Representations, Assessed via Job Transition Data"(Poster) | Tyler Baldwin, Wyatt Clarke, Maysa Macedo, Rogerio Paula, and Subhro Das |
| --- | --- |
| N270 "Data Utilization in the Education Field by Local and Central Governments in Japan"(Poster) | Hikojiro Fujiwara and Satoru Hagino |

| | |
|---|---|
| N284 "Decomposing User-APP Graph into Subgraphs for Effective APP and User Embedding Learning"(Poster) | Tan Yu, Jun Zhi, Yufei Zhang, Jian Li, Hongliang Fei, and Ping Li |

## I&G S11: Big Data Analysis (2)

| | |
|---|---|
| N286 "Scalable Household Identification using Mobile Engagement Data - A Weighted Two-Mode Network Approach"(Regular) | Vishnu Gowthem Thangaraj, Kajanan Sangaralingam, Nisha Verma, and Anindya Dutta |
| N241 "Mining Railway Grievances on Twitter for Efficient E-Governance in India"(Regular) | Swati Agarwal, Ashrut Kumar, and Rijul Ganguly |
| N267 "Implementation of Biased Big Data to the Japanese Official Labor Statistics Using Supervised Learning under Covariate Shift"(Regular) | Yuya Takada and Kiyoshi Izumi |
| N249 "Predictive model of frailty onset using Bayesian network"(Regular) | Yujiro Kawai, Eiichi Sakurai, Yuki Sugawara, Yukihiko Okada, Kai Tanabe, Akiko Tsukao, and Shinya Kuno |
| N250 "Earth Observation Data Utilization for SDGs Indicators: 15.4.2 and 11.3.1"(Poster) | Osamu Ochiai, Mariko Harada, and Ko Hamamoto |

## I&G S12: Big Data Systems (2)

| | |
|---|---|
| N262 "StreamFlow: A System for Summarizing and Learning Over Industrial Big Data Streams"(Short) | Mariam BARRY, Saad EL JAOUHARI, Albert BIFET, Eric GUERIZEC, Jacob MONTIEL, and Raja CHIKY |
| N247 "HAIVAN: a Holistic ML Analytics Infrastructure for a Variety of Radio Access Networks"(Poster) | Hong-Linh Truong and Nguyen Ngoc Nhu Trang |
| N291 "FeatureBox: Feature Engineering on GPUs for Massive-Scale Ads Systems"(Poster) | Weijie Zhao, Xuewu Jiao, Xinsheng Luo, Jingxue Li, Belhal Karimi, and Ping Li |
| N272 "Automating privacy compliance in the decentralized enterprise"(Poster) | Gero Dittmann, Chris Giblin, Michael Osborne, and Rahul |
| N222 "Representation Learning on Graphs to Identifying Circular Trading in Goods and Services Tax"(Poster) | Priya Mehta, Sanat Bhargava, Sandeep Kumar, Ravi Kumar, and Sobhan Babu Ch |

## I&G S13: Big Data Applications (1)

| | |
|---|---|
| N235 "Gradual Further Pre-training Architecture for Economics/Finance Domain Adaptation of Language Model"(Poster) | Hiroki Sakaji, Masahiro Suzuki, Kiyoshi Izumi, and Hiroyuki Mitsugi |
| N223 "SSAAM: Sentiment Signal-based Asset Allocation Method with Causality Information"(Poster) | Rei Taguchi, Hiroki Sakaji, and Kiyoshi Izumi |
| N273 "Detecting Group Behavior for Anti-Money Laundering With Incomplete Network Information"(Poster) | Pavlo Tertychnyi, Tommy Lindström, Changling Liu, and Marlon Dumas |

## I&G S14: Big Data Analysis (3)

| | |
|---|---|
| N271 "Big data techniques for Industrial problem with little data"(Poster) | Elham Khabiri, Bhavna Agrawal, Joseph Lindquist, Yingjie Li, and Anuradha Bhamidipaty |
| N207 "TgraphSpot: Fast and Effective Anomaly Detection for Time-Evolving Graphs"(Poster) | Mirela Cazzolato, Saranya Vijayakumar, Xinyi Zheng, Meng-Chieh Lee, Namyong Park, Pedro Fidalgo, Bruno Lages, Agma Traina, and Christos Faloutsos |

## I&G S15: Big Data Applications (2)

| | |
|---|---|
| N215 "Decision Support for Disability Employment using Counterfactual Survival Analysis"(Regular) | Ha Tran, Thuc Le, Jiuyong Li, Lin Liu, Jixue Liu, Yanchang Zhao, and Tony Waters |
| N238 "System and Method on Order Management Using Neural Networks and Risk Modeling" (Regular) | Shubhi Asthana, Bing Zhang, Pawan Chowdhary, and Taiga Nakamura |

| N254 "Sensing Multi-modal Mobility Patterns: A Case Study of Helsinki using Bluetooth Beacons and a Mobile Application" (Regular) | Zhiren Huang, Alonso Espinosa, and Charalampos Sipetas |
|---|---|
| N225 "Similarity-based Historical Input Selection to Predict Irregular Holiday Traffics in Real-time"(Short) | Pooreumoe Kim, Jungsoo Park, Seungyoun Lee, and Jungmin Kim |
| N228 "The Future of Extreme Weather Events - Advanced Machine Learning and Artificial Intelligence for Democratic Institution preparedness and enhanced National Food Security!" | Chandrasekar Vuppalapati, Anitha Ilapakurti, Sharat Kedari, Jaya Shankar Vuppalapati, and Santosh Kedari |
| N264 "Implementation of Empath X SLA predictive tool for a Government Agency"(Short) | Alvina Hui Shan Lee, Venky Shankararaman, and Eng Lieh Ouh |
| **I&G S16: Big Data Analysis (4)** | |
| N295 "Flexible Partitioning of Geographical Information based on GPS Coordinates for k-Anonymity"(Regular) | Liu Yuxin and Kazuhiro Minami |
| N258 "A practical facility location optimization with uncertain variables in emergency road services"(Regular) | Shogo Takedomi and Tsukasa Ishigaki |
| N276 "Matching Attacks on Non-deterministic Algorithms for Cell Suppression Problem for Tabular Data"(Short) | Yutaka Abe and Kazuhiro Minami |
| N245 "Explainable Machine Learning Models for Evaluating Government Grantmaking"(Short) | Daniel Fonner and Frank Coyle |
| N259 "Energy-based Domain Adaption with Active Learning for Emerging Misinformation Detection"(Poster) | Kyumin Lee, Guanyi Mou, and Scott Sievert |
| N248 "Analysis and Estimation of News Article Reading Time with Multimodal Machine Learning"(Poster) | Shotaro Ishihara and Yasufumi Nakama |

# Workshops:

| **Computational Archival Science**<br>*Workshop Chairs: Mark Hedges, Victoria Lemieux, Richard Marciano*<br>***Monday 19 December (all times are in Mountain Standard Time, UTC-7)*** | | |
|---|---|---|
| **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 | Giovanni Colavizza (U. Amsterdam) |
| **10:00 – 11:00** | **SESSION 1: Computational approaches in archival practice** | |
| 10:00 – 10:20 | #1: An Intelligent Class: The Development Of A Novel Context Capturing Method For The Functional Auto Classification Of Records (S01207) | N. Payne |
| 10:20 – 10:40 | #2: A Data-Driven Approach to Reparative Description at the University of Chicago (S01211) | A. Gosselar |

| | | |
|---|---|---|
| 10:40 – 11:00 | #3: Metadata Verification: A Workflow for Computational Archival Science (S01216) | J. Pepper, A. Senin, D. Jebbia, D. Breen, J. Greenberg |
| 11:00 – 11:20 | Coffee Break (coffee not supplied!) | |
| 11:20 – 12:40 | SESSION 2: Working with archival materials | |
| 11:20 – 11:40 | #4: AI and archive. Handwritten Text Recognition Applied to Patrimonial Holdings: An Example of 10 diaries written by Spanish Republican Teachers in 1932 (S01214) | P. Raventos, C. Hernandez, M. Simon |
| 11:40 – 12:00 | #5: CensusIRL: Historical census data preparation with MDD support (S01213) | L. A. Doherty, R. Murphy, A. Schieweck, S. Clancy, C. Breathnach, T. Margaria |
| 12:00 – 12:20 | #6: The Arquive of Tatuoca Magnetic Observatory Brazil: from paper to intelligent bytes (S01209) | C. Berrío-Zapata, E. F. da Silva, M. C. Pinheiro, C. M. Martins, V. A. C. de Abreu, M. A. Góngora, K. Dunman |
| 12:20 – 12:40 | #7: Applications of data analysis on scholarly long documents (S01212) | B. Banerjee, W. A. Ingram, J. Wu, E. A. Fox |
| 12:40 – 1:20 | Discussion | |
| 1:20 – 2:00 | Lunch Break (lunch not supplied!) | |
| 2:00 – 2:40 | SESSION 3: CAS and education | |
| 2:00 – 2:20 | #8: Computational Thinking Integration into Archival Educators' Networked Instruction (S01215) | S. Buchanan, K. Gracy, J. Kitchens, R. Marciano |
| 2:20 – 2:40 | #9: Teachable Insights from Working with The Mary Eliza Project for Gastronomy Students Working with Data (S01205) | L. Kitchings |
| 2:40 – 3:20 | SESSION 4: New forms of archives | |
| 2:40 – 3:00 | #10: A Technical Assessment of Blockchain in Healthcare with a Focus on Big Data (S01206) | G. Al-Sumaidaee, A. Alexandridis, R. Alkhudary, Z. Zilic |
| 3:00 – 3:20 | #11: Evolutionary Archives: The Unlikely Comparison of GenBank and Know Your Meme (S01217) | S. Bratt, A. O. Smith |
| 3:20 – 3:40 | Coffee Break (coffee not supplied!) | |
| 3:40 – 4:20 | SESSION 5: Preservation and data management | |
| 3:40 – 4:00 | #12: LABDRIVE, a Petabyte scalable, OAIS/ISO 16363 conformant, for scientific research organisations to preserve documents, processed data, and software (S01210) | D. Giaretta, T. Redondo, A. G. Martinez, M. Fuertes |
| 4:00 – 4:20 | #13: Open Science and Research Data Management: A FAIR European Postgraduate Programme (S01208) | H. Gonzalez-Velez, C. Dobre, B. Sanchez-Solis, G. Antinucci, D, Feenan, D. Gheorghe |
| 4:20 – 4:40 | SESSION 6: Working with archival materials (continued) | |
| 4:20 – 4:40 | #14: Event Time Extraction from Japanese News Archives (S01204) | S. Peng, A. Yamamoto, S. Mori, T. Sekino, |
| 4:40 – 5:20 | SESSION 5: Discussion and closing | |

## Workshop on Cyber Threat Intelligence and Hunting (CyberHunt 2022)

**Workshop co-chairs:** Vasileios Mavroeidis and Mateusz Zych
**Program committee chair:** Andre Büttner

| Time | Title | Presenter/Author |
|---|---|---|
| 14:00-14:05 | **Opening Remarks** | Vasileios Mavroeidis |
| 14:05-14:35 | Keynote Presentation: Towards Integrated and Adaptive Cybersecurity Operations with Cross-Border Collaboration - A European Perspective | Konstantinos Fysarakis |
| 14:35-15:05 | FeedRef2022: A Named Entity Recognition Dataset for Extracting Indicators of Compromise (S02203) | Hsin-Ju Chan |
| 15:05-15:35 | An Effect Analysis of ISO/IEC 27001 Certification on Technical Security of Norwegian Grid Operators (S02205) | Øyvind Toftegaard |
| 15:35-16:05 | Towards XAI in the SOC - a user centric study of explainable alerts with SHAP and LIME (S02207) | Gudmund Grov |
| 16:05-16:30 | **Coffee Break** | |
| 16:30-17:00 | LADEMU: a modular & continuous approach for generating labelled APT datasets from emulations (S02208) | Gudmund Grov |
| 17:00-17:30 | A tool for assisting in the forensic investigation of cyber-security (S02212) | Konstantinos Touloumis |
| 17:30-18:00 | Towards Dynamic Capture-The-Flag Training Environments For Reinforcement Learning Offensive Security Agents (S02213) | Robert Andrew Chetwyn |
| 18:00-18:30 | A Blueprint for Collaborative Cybersecurity Operations Centres with Capacity for Shared Situational Awareness, Coordinated Response, and Joint Preparedness (S02214) | Konstantinos Fysarakis |
| 18:30-18:40 | **Closing Remarks** | Vasileios Mavroeidis |

| 8th International Workshop on Methods to Improve Big Data Science Projects December 18th (14:00 to 18:30 Osaka Time) *WorkshopChair: Jeffrey Saltz* | | |
|---|---|---|
| **Time** | **Title** | **Presenter/Author** |
| 14:00 | *S03207: An Organizational Framework of Institutional Stakeholder Engagement for Capacity to Support Big Data Science Teams Towards Cyberinfrastructure Diffusion* | Kerk Kee |
| 14:20 | *S03205: Anovos: A Scalable Feature Engineering Library* | Kajanan Sangaralingam |
| 14:40 | *S03203: Project Artifacts for the Data Science Lifecycle: A Comprehensive Overview* | Christian Haertel |
| 15:00 | *S03206: A Framework for the Design, Development, Testing and Deployment of Reliable Big Data Platforms* | Catherine Inibhunu |
| 15:20 | *BigD270: A Novel Rigorous Measurement Model for Big Data Quality Characteristics* | Haochen Zou |
| **15:40** | **Coffee Break** | |
| 15:50 | *S03209: Adaptive Method for Machine Learning Model Selection in Data Science Projects* | Cristina Tavares |
| 16:10 | *BigD670: A Labeling Task Design for Supporting Recent Algorithmic Needs* | Jaeyoun You |
| 16:30 | *S03204: Nine Questions to Evaluate a Data Science Team's Process: Exploring a Big Data Science Team Process Evaluation Framework Via a Delphi Study* | Jefrey Saltz |

| | | |
|---|---|---|
| 16:50 | *S03208: Analyzig Data Science Online Practitioner Communities: Trends and Implications for Data Science Project Management* | Sucheta Lahiri |
| **17:10** | **Closing Remarks** | |

## Workshop on Big Data Analytic in Healthcare Tentative Schedule
### (8:30am-12:30pm, Dec 17, ONLINE)
Workshop Chairs: Alex Kuo & Elizabeth Borycki

| Time | Title | Presenter/Author |
|---|---|---|
| 8:40-9:00am, Dec 17 | S05202: *Usability Evaluation of a Mobile Emotion Healthcare System Applying Sentiment analysis* | Shu-Lin Wang |
| 9:00-9:20am, Dec 17 | S05203: *Big data analytics of predicting annual US Medicare billing claims with health services* | Dillon Chrimes |
| 9:20-9:40am, Dec 17 | S05205: *Text Generation to Aid Depression Detection: A Comparative Study of Conditional Sequence Generative Adversarial Networks* | ML Tlachac |
| 9:40-10:00am, Dec 17 | S05208: *The Dimensional Analysis of Data Flow Programs That Include Multidimensional and User-Defined Functions* | Monem Shennat |
| 10:00-10:20am, Dec 17 | S05209: *A Pre-trained Multi-modal Model for Understanding Scanned Document in Electronic Health Records* | Qiang Wei |
| 10:20-10:40am, Dec 17 | BigD220: *A Data-driven Latent Semantic Analysis for Automatic Text Summarization using LDA Topic Modelling* | Daniel F.O. Onah |
| 10:40-11:00am, Dec 17 | BigD584: *Mining Latent Disease Factors from Medical Literature using Causality* | Pranav Gujarathi |
| 11:00-11:20am, Dec 17 | BigD602: *Deep Learning Models with Stratification-based Loss Function on Domain Knowledge-based Time series Data : Hypotension Prediction* | Hanseok Jeong |
| 11:20-11:40am, Dec 17 | BigD697: *Constructing Machine Learning models for Orthodontic Treatment Planning; a comparison of different methods* | Hamid Shojaei |
| 11:40-12:00am, Dec 17 | BigD720: *End-to-end Multiple Instance Learning with Gradient Accumulation* | Axel Andersson |
| 12:00-12:20pm, Dec 17 | S05210: *Classification of Heart Sound using Tunable Q Wavelet Transform and Machine Learning* | Ying Xie |
| 12:20-12:30pm, Dec 17 | **Closing Remarks: Alex Kuo (workshop chair)** | |

## Workshop Title: Big Data Engineering and Analytics in Cyber-Physical Systems (BigEACPS'22)
*Workshop Chair: Akbar Namin (December 19, 2022 – Afternoon)*

| Time | Title | Presenter/Author |
|---|---|---|
| 14:00 – 14:25 | *(Paper ID: S06203) Anomaly Detection for Public Transport and Air Pollution Analysis* | Paolo Mignone, Donato Malerba, and Michelangelo Ceci |
| 14:25 – 14:50 | *(Paper ID: S06206) Solar Irradiance Prediction Using Transformer Machine Learning Models* | Ayda Demir, Luis Felipe Gutiérrez, Akbar Siami Namin, and Stephen Bayne, |
| 14:50 – 15:15 | *(Paper ID: S06207) Assessing Energy Consumption in Data Acquisition from Smart Wearable Sensors in IoT-Based Health Applications* | Antonio Paganelli, André Sarmento, Adriano Branco, Markus Endler, Nathalia Nascimento, Paulo Alencar, and Donald Cowan |
| 15:15 – 15:40 | *(Paper ID: S06208) SIMCal: A High-Performance Toolkit For Calibrating Traffic Simulation* | Toan Tran, Seyedmehdi Khaleghian, Junxuan Zhao, and Mina Sartipi |

| | | |
|---|---|---|
| 15:40 – 16:05 | *(Paper ID: S06201) Metaverse in Education: Vision, Opportunities, and Challenges* | Hong Lin, Shicheng Wan, Wensheng Gan, Jiahui Chen, and Han-Chieh Chao |
| 15 min | **Break** | |
| 16:20 – 16:45 | *(Paper ID: S06210) Classifying Perceived Emotions based on Polarity of Arousal and Valence from Sound Events* | Pooja Krishan and Faranak Abri |
| 16:45 – 17:10 | *(Paper ID: S06211) Himanshu Neema, and Abhishek Dubey, BTE-Sim: Fast simulation environment for public transportation* | Rishav Sen, Toan Tran, Seyedmehdi Khaleghian, Philip Pugliese, Mina Sartipi |
| 17:10 – 17:35 | *(Paper ID: S06202) FedProb: An Aggregation Method Based on Feature Probability Distribution for Federated Learning on Non-IID Data* | Do-Van Nguyen, Anh-Khoa Tran, and Koji Zettsu |
| 17:35 – 18:00 | *(Paper ID: S06204) Using Transformers for Identification of Persuasion Principles in Phishing Emails* | Bimal Karki, Faranak Abri, Akbar Namin, and Keith Jones |
| 18:00 – 18:20 | *(Paper ID: S06209) A Study On Security and Privacy Risks of Self-Disclosure On Social Networking Sites During COVID-19 Pandemic* | Oluwabunmi Ariyo and Jianjun Zheng |
| | **Closing Remarks** | |

## GTA³ 6.0: The 6th workshop on Graph Techniques for Adversarial Activity Analytics
*WorkshopChairs: Jiejun Xu, Hanghang Tong, Andrea Bertozzi, Vince Lyzinski, George Chin, Hang Liu, Tyler Derr*

| Time (JST) | Title | Presenter/Author |
|---|---|---|
| 8:30 am | **Opening Remarks** | |
| 8:35 – 9:20 am | **Keynote 1** | Prof. Wei Wang |
| 9:20 – 10:05 am | **Keynote 2** | Prof: Ross Maciejewski |
| 10:05 – 10:25 am | Uncertainty Visualization for Graph Coarsening (S07201) | Fangfei Lan, Sourabh Palande, Michael Young, and Bei Wang |
| 10:25 – 10:35 am | **Coffee Break** | |
| 10:35 – 11:20 am | **Keynote 3** | Dr. Aydin Buluc |
| 11:20 – 11:40 am | Dual Graphs of Polyhedral Decompositions for the Detection of Adversarial Attacks (S07202) | Huma Jamil, Yajing Liu, Christina Cole, Nathaniel Blanchard, Emily King, Michael Kirby, and Christopher Peterson |
| 11:40 – 12:00 pm | Knowledge Graphs of the QAnon Twitter Network (S07203) | Clay Adams, Malvina Bozhidarova, James Chen, Andrew Gao, Zhengtong Liu, Hunter Priniski, Junyuan Lin, Rishi Sonthalia, Andrea Bertozzi, and Jeffrey Brantingham |
| 12:00 – 12:20 pm | A Survey of Explainable Graph Neural Networks for Cyber Malware Analysis (S07204) | Dana Warmsley, Alex Waagen, Jiejun Xu, Zhining Liu, and Hanghang Tong |
| | | |
| 12:20 pm | **Closing Remarks** | |

## 5th Workshop on Big Data for CyberSecurity (BigCyber-2022)
*WorkshopChairs: Dr. Karuna Joshi, Dr. Rajeev Agrawal, Dr. Sudip Mittal*

| Time (JST) | Title | Presenter/Author |
|---|---|---|
| 14:00 - 14:10 | *Introductions* | |
| 14:10 – 14:30 | *Unsupervised Learning Approaches for Construction of Malware Families (S08209)* | Alexis Jackson and Krishnendu Ghosh |
| 14:30 – 14:50 | *The Lannion report on Big Data and Security Monitoring Research (S08204)* | Laurent d'Orazio, Juba Agoun, Elisa Bertino, Georges Bossert, Jalil Boukhobza, Le Gruenwald, Mohand-Saïd Hacid, Van Long Nguyen Huu, Makoto Onizuka, Omer Rana, Hervé Rannou, Taofik Saïdi, and Dimitri Tombroff |
| 14:50 – 15:10 | *Detecting Malware Activity Using Public Search Data (S08210)* | Ismael Villanueva Miranda and Monika Akbar |
| 15:10 – 15:30 | *Benchmark of Data Preprocessing Methods for Imbalanced Classification (S08207)* | Radovan Haluška, Jan Brabec, and Tomáš Komárek |
| 15:30 – 15:50 | *Saisiyat Is Where It Is At! Insights Into Backdoors And Debiasing Of Cross Lingual Transformers For Named Entity Recognition (BigD367)* | Ricardo Calix, JJ Ben-Joseph, Nina Lopatina, Ryan Ashley, Mona Gogia, George Sieniawski, and Andrea Brennen |
| 15:50 – 16:10 | *Break* | |
| 16:10 – 16:30 | *Critical-Threat-Alert Detection using Online Learning (S08205)* | Samuel Ndichu, Tao Ban, Takeshi Takahashi, and Daisuke Inoue |
| 16:30 – 16:50 | *Metaverse Security and Privacy: An Overview (S08203)* | Zefeng Chen, Jiayang Wu, Wensheng Gan, and Zhenlian Qi |
| 16:50 – 17:10 | *Dark patterns in e-commerce: a dataset and its baseline evaluations (S08202)* | Yuki Yada, Jiaying Feng, Tsuneo Matsumoto, Nao Fukushima, Fuyuko Kido, and Hayato Yamana |
| 17:10 – 17:30 | *Darknet Analysis-Based Early Detection Framework for Malware Activity: Issue and Potential Extension (S08201)* | Chansu Han, Akira Tanaka, and Takeshi Takahashi |
| 17:30 – 17:50 | *Detecting registrants anomalies for .it ccTLD (BigD382)* | Arianna Del Soldato, Maurizio Martinelli, and Daniele Sartiano |
| 17:50 – 18:00 | *Closing Remarks* | |

## The 6th International Workshop on Big Data Analytic for Cyber Crime Investigation and Prevention
Saturday, December 17th, 2022; Location: hybrid (GMT+9, Osaka, Japan)
Workshop Chairs: ***Andrii Shalaginov, Asif Iqbal, Igor Kotsiuba, Mamoun Alazab***
Kristiania University College, Norway
KTH Royal Institute of Technology in Stockholm, Sweden
Durham University, UK
Charles Darwin University, Australia

| Time (GMT +9) | Time (CET) | Title of the talk | Presenter/Author |
|---|---|---|---|
| 09:00-09:05 | 01:00-01:05 | **Opening Remarks and Welcome** | Andrii Shalaginov |
| 09:05-09:45 | 01:05-01:45 | **Cybersecurity in Europe - A police research perspective: S4ALLCITIES and CC-DRIVER projects** | Keynote speaker: **José L. Diego, Valencia Local Police, Spain** |
| 09:45-10:10 | 01:45-02:10 | *S09202: On the Potential of Using ERP Business and System Data for Fraud Detection* | Jonas Schnepf, Paula Vetter, Tarik Temel, Bernd Scheuermann, and Lars Schmidt-Thieme |
| 10:10-10:35 | 02:10-02:35 | *S09203:Evaluating Anti-Virus Effectiveness in Linux* | Giuseppe Raffa, Daniele Sgandurra, and |

| | | | Daniel O'Keeffe |
|---|---|---|---|
| 10:35-11:00 | 02:35-03:00 | *S09204: DBNex: Deep Belief Network and Explainable AI based Financial Fraud Detection* | Abhimanyu Bhowmik, Madhushree Sannigrahi, Deepraj Chowdhury, Ashutosh Dhar Dwivedi, and Raghava Rao Mukkamala |
| **11:00-11:15** | **03:00-03:15** | | ***Coffee Break*** |
| 11:15-11:40 | 03:15-03:40 | *S09207: Investigating User Information and Social Media Features in Cyberbullying Detection* | Jiabao Qiu, Nihar Hegde, Melody Moh, and Teng-Sheng Moh |
| 11:40-12:05 | 03:40-04:05 | *S09209: Threat Miner - A Text Analysis Engine for Threat Identification Using Dark Web Data* | Nathan Deguara, Junaid Arshad, Anum Paracha, and Muhammad Ajmal Azad |
| 12:05-12:30 | 04:05-04:30 | *S09210: Secure Application of Artificial Intelligence Applications for Anti-Money Laundering using Confidential Computing* | Richard Searle, Prabhanjan Gururaj, Anubhav Gupta, and Kiran Kannur |
| 12:30-12:55 | 04:30-04:55 | *S09211:Artificial Intelligence Enabled Middleware for Distributed Cyberattacks Detection in IoT-based Smart Environments* | Guru Prasad Bhandari, Andreas Lyth, Andrii Shalaginov, and Tor-Morten Grønli |
| **12:55-13:10** | **04:55-05:10** | | ***Coffee Break*** |
| 13:10-13:35 | 05:10-05:35 | ***S09212: Cyber Security of Electric Vehicle Charging Infrastructure: Open Issues and Recommendations*** | **Inna Skarga-Bandurova, Igor Kotsiuba, and Tetiana Biloborodova** |
| 13:35-14:00 | 05:35-06:00 | *BigD384: Forensic Analysis of Amazon Alexa Echo Dot 4th Generation* | Damilola Oladimeji and Bing Zhou |
| 14:00-14:05 | 06:00-06:05 | **Closing Remarks** | Andrii Shalaginov |

| The 2nd International Workshop on Multi-Modal Medical Data Analysis (The 2nd International Workshop on Multi-Modal Medical Data Analysis)*WorkshopChairs: Bo Zhang* *zbo@bupt.edu.cn* | | |
|---|---|---|
| **Time** | **Title** | **Presenter/Author** |
| 2:00 pm – 2:15 pm | *Improving the performance of disease detection from chest X-ray images using deep learning* | Duong Hong Nguyen |
| 2:15 pm – 2:30 pm | *BioSentEval: An Evaluation Toolkit for Chinese Medical Sentence Representation* | Mingjun Gu |
| 2:30 pm – 2:45 pm | *ACL-Net: Adaptive and Collaborative Learning Network for Multi-Site Prostate MRI Segmentation* | Zibo Ma |
| 2:45 pm – 3:00 pm | *Enabling Scalable Analytics of Physiological Sensor and Derived Feature Multi-Modal Time-Series with Big Data Management* | Jamal Saied-Walker |
| 3:00 pm – 2:15 pm | *Classification of Autism Spectrum Disorder Using rs-fMRI data and Graph Convolutional Networks* | Tianren Yang |
| | **Closing Remarks** | |

| **The 3rd International Workshop on Big Data Reduction (IWBDR-3)** *WorkshopChairs: Dingwen Tao, Xin Liang, Sheng Di* *dingwen.tao@wsu.edu* | | |
|---|---|---|
| **Time** | **Title** | **Presenter/Author** |
| 08:00 – 08:05 | Opening Remarks | Dingwen Tao, Xin Liang, Sheng Di |
| 08:05 – 09:00 | Keynote: TBD | TBD |
| 09:00 – 09:20 | S11201: Machine Learning Platform for Extreme Scale Computing on Compressed IoT Data | Seshu Tirupathi |
| 09:20 – 09:40 | S11202: Lossy Compression to Reduce Latency of Local Image Transfer for Autonomous Off-Road Perception Systems | Max Faykus |
| 09:40 – 10:00 | BigD616: Estimating Potential Error in Sampling Interpolation | Megan Hickman |
| 10:10 – 10:20 | **Coffee Break** | |

| Time | Title | Presenter/Author |
|---|---|---|
| 10:20 – 10:40 | S11204: Towards Guaranteeing Error Bound in DCT-based Lossy Compression | Jiaxi Chen |
| 10:40 – 11:00 | S11203: Exploring Data Corruption Inside SZ | Ruiwen Shan |
| 11:00 – 11:20 | BigD247: Extraction of Power Consumption Patterns using Non-negative Tucker Decomposition | Taku Moriyama |
| **11:20 – 11:30** | **Closing Remarks** | |

## 10th International Workshop on Distributed Storage and Blockchain Technologies for Big Data
*WorkshopChairs: Hui Li and Han Wang*

| Time | Title | Presenter/Author |
|---|---|---|
| 10:30 - 11:00 | S12201, Some Results on Optimal Octonion-Field Locally Repairable Codes Of Distance 3 and 4 | Jieying Liu, Guojun Han, and Hanxu Hou |
| 11:00 - 11:30 | S12202, Systematic MDS Array Codes Correcting a Single Criss-Cross Error with Lower Update Complexity | Lin Zou, Hanxu Hou, and Xiaoan Zhou |
| 11:30 - 12:00 | S12204, Optimizing Repair-Cost of Locally Repairable Codes for Hot Data in Cluster Storage Systems | Lu Zhong, Goujun Han, Hanxu Hou, and Qingda Hai |
| 12:00 - 12:30 | S12205, Blockchain Interoperability Landscape | Inwon Kang, Aparna Gupta, and Oshani Seneviratne |
| 12:30 - 14:00 | **Lunch Break** | |
| 14:00 – 14:30 | S12207, Optimizing Parallel Proof of Vote Consensus Based on Mimic Security in Consortium Blockchains | Zhenwei Xiao, Hui Li, Han Wang, Guoliang Yang, Qiongwei Ye, Shusheng Zou, Ping Lu, and Qi Lyu |
| 14:30 – 15:00 | S12208, An Improved Vulnerability Detection System of Smart Contracts Based on Symbolic Execution | Yao Yao, Hui Li, Xin Yang, and Yiwang Le |
| 15:00 – 15:30 | BigD541, A Simple Approach to Optimize S3 Object Gateways for Massive Numbers of Small File Writes | Patrick Cato |
| 15:30 – 16:00 | BigD740, Rental NFT Protocol With Advanced Rewards Splitting | Andrew Kirillov, Abrar Rahman, and Ayush Aggarwal |
| 16:00 – 16:30 | **Coffee Break / Keynote Speech** | |

## New Strands of Usage of Big Data in Medical Systems, Market and Institutional Design, and Economic Theory (WS 13) In-Person (partly online)
*Workshop Chair: Naoki Watanabe*
*naoki50@keio.jp*

| Time | Title | Presenter / Co-Author(s) |
|---|---|---|
| 14:00-14:05 | Organizer's Notes | Naoki Watanabe |
| 14:05-14:35 | Real World Data Analysis in Health Technology Assessment (online) | Rei Goto (invited speaker) |
| 14:35-15:00 | (S13204_3622) Predicting the Impact of Shared Autonomous Vehicle Service on Tokyo Transportation Using MATSim | Yutaka Ishibashi / Eizo Akiyama |
| 15:00-15:25 | (S13202_7543) Land Consolidation by Plot Exchange | Kengo Kurosaka / Naoki Onodera |
| 15:25-15:35 | **Coffee Break** | |
| 15:35-16:00 | (S13201_8815) Feedback Information on Cumulative Payoff in a Bandit Experiment: Meaningful Learning in Weighted Voting | Kazuhito Ogawa / Naoki Watanabe |
| 16:00-16:25 | (S13206_5596) Network Formation of Next Generation Electric Power and Mobility in Rural Area | Yumi Watahiki / Hiroyasu Ando, Isamu Takahara |
| 16:25-16:50 | (S13203_8537) A Model of Pricing Data and Their Constituent Variables Traded in Two-Sided Markets with Resale; A Subject Experiment | Toshihiko Nanba / Kazuhito Ogawa, Naoki Watanabe, Teruaki Hayashi, Hiroki Sakaji |
| 16:50-17:00 | **Coffee Break** | |
| 17:00-17:15 | (S13207_4286) Compatibility of Efficiency and Stability in Priority-Augmented Object Allocation Problems with Weak Priorities (short talk) | Yusuke Iwase / Wataru Ishida |

| | | |
|---|---|---|
| 17:15-17:30 | (S13205_6043) Designing a Mechanism to Promote the Smooth Provision of Medical Data for the Utilization of Artificial Intelligence in Life Science (short talk) | Masanori Fujita / Ayato Kitadai, Koichi Sumikura, Nariaki Nishino |
| 17:30-17:55 | (S13208_5552) A Subject Experiment of an Approximate DGS Algorithm: Price Increment, Allocative Efficiency, and Sellers's Revenue | Naoki Watanabe / Yoichi Izunaga, Satoshi Takahashi |
| 17:55-18:00 | **Closing Remarks** | |

## The Sixth Annual Workshop on Applications of Artificial Intelligence in the Legal Industry
*Workshop Chairs: Jianping Zhang, Nathaniel Huber-Fliflet, Robert Keeling, Christian J. Mahoney, Han Qin, Adam Dabrowski*
*zhaohaozhen@gmail.com,jianping.zhang@ankura.com,nathaniel.huber-fliflet@ankura.com*
https://ankura.zoom.us/webinar/register/WN_WdXc7S_FRfaeUWXSViGKSg

| Time | Title | Presenter/Authors |
|---|---|---|
| 9:30-9:40 am | **Opening Remarks** | |
| 9:40-10:00 am | Exploring Technology Assisted Review Methodologies for Multilingual Data Sets | Christian J. Mahoney, Chris T. Clark, Qiang Mao, Fusheng Wei |
| 10:00:10:20 am | Extracting Crime Prosecution Elements based on Neural Machine Reading Comprehension Model (Paper ID: S14204) | Jui-Ching Tsou, Kai-Yu Hsieh, Chen-Hua Huang, Yu-An Shih, Han-Cheng Yu, and Yao-Chung Fan |
| 10:20-10:40 am | An Empirical Comparison of DistilBERT, Longformer and Logistic Regression for Predictive Coding (Paper ID: S14205) | Fusheng Wei, Jingchao Yang, Qiang Mao, Han Qin, and Adam Dabrowski |
| 10:40-11:10 am | **Invited Talk** | Christian J. Mahoney |
| 11:10-11:40 am | **Coffee Break** | |
| 11:40-12:00 am | Data protections for minors with named entity recognition (Paper ID: S14206) | Jason Gillette, Sayed Khushal Shah, Zeenat Tariq, and Shabbab Algamdi |
| 12:00-12:30 pm | **Invited Talk** | TBD |
| 12:30-12:50 am | An Application of Object Detection in Contract Analysis (Paper ID: S14201) | Josh Levy, Fusheng Wei, Jingchao Yang, and Adam Dabrowski |
| 12:50-13:10 am | A Business Workflow For Providing Open-Domain Question Answering Reader Systems on The Wikipedia Dataset (Paper ID: S14203) | Dilan Bakır and Mehmet S. Aktas |
| 13:10-13:20 pm | **Closing Remarks** | |

## Workshop Ti7th Workshop on Real-time Stream Analytics, Stream Mining, CER/CEP & Stream Data Management in Big Data
*WorkshopChairs: Sabri SKHIRI*
*sabri.skhiri@euranova.eu*

| Time | Title | Presenter/Author |
|---|---|---|
| 2:00 PM - 2:40 PM | Keynote 1: SQL Extensions To Support Streaming Data | Fabian Hueske, Snowflake |
| 2:40 - 3:20 | Keynote 2: A story of batch and stream processing at Google and elsewhere | Pablo Estrada, Google |
| 3:20- 4:00 | Keynote 3: Govern your streams through a data products foundation | Marc Delbaere, Digazu |
| 4:00 - 4:20 | Dynamic Ensemble Size Adjustment for Memory Constrained Mondrian Forest | Martin Khannouz and Tristan Glatard |
| 4:20 - 4:30 | **Coffee Break** | |
| 4:30: 4:50 | Effective Weighted k-Nearest Neighbors for Dynamic Data Streams. | Maroua Bahri, |
| 4:50 - 5:10 | Prequential Model Selection for Time Series Forecasting based on Saliency Maps | Shivani Tomar, seshu Tirupathi, Dhaval Vinodbhai Salwala, Ivana Dusparic, Elizabeth Daly |
| 5:10 - 5:30 | An LSTM Encoder-Decoder Approach for Unsupervised Online Anomaly Detection in Machine Learning Packages for Streaming Data. | Nabil Belacel, Rene Richard, Zhicheng Xu |
| 5:30 - 5:50 | Are Concept Drift Detectors Reliable Alarming Systems? - A Comparative Study. | Lorena Poenaru-Olaru, Luis Cruz, Arie van Deursen, Jan Rellermeyer |

| Time | Title | Presenter/Author |
|---|---|---|
| 5:50 - 6:10 | Region-based Sub-Snapshot (RegSnap): Enhanced Fault Tolerance in Distributed Stream Processing with Partial Snapshot. | Takdir Takdir, Hiroyuki Kitagawa. |
| **6:10 - 6:15** | **Closing Remarks** | |

## The 4th International Workshop on Big Data Tools, Methods, and Use Cases for Innovative Scientific Discovery (BTSD) 2022

*Workshop Chairs: Matt Sangkeun Lee, Anika Tabassum, Jong Youl Choi – December 17, 2022 (Virtual Workshop)*

| Time | Title | Presenter/Author |
|---|---|---|
| 8:30am-8:40am | **Opening Remarks** | |
| 8:40am-9:00am | Improving Predictions Under Uncertainty of Material Plasma Device Operations (Short) - S16201 | Rick Archibald <archibaldrk@ornl.gov> |
| 9:00am-9:25am | Adaptive Generation of Training Data for ML Reduced Model Creation (Long) - S16203 | Mark Ciancosa <cianciosamr@ornl.gov> |
| 9:25am-9:45am | Real-time Multi-granular Analytics Framework for HIT Systems (Short) - S16205 | Byung Park <parkbh@ornl.gov> |
| 9:45am-10:10am | Partial Discharge Detection for Underground Transmission Lines Using Nonnegative Matrix Factorization (Long) - S16204 | Akihiro Tanabe <atanabe88@sanken.osaka-u.ac.jp> |
| 10:10am-10:30am | Collaboration in relation to Human-AI Systems: Status, Trends, and Impact (Short) - S16207 | Antonio Correia <antonio.g.correia@inesctec.pt> |
| 10:30am-10:40am | **Coffee Break & Discussion** | |
| 10:40am-11:00am | Distilling Knowledge from Ensembles of Cluster-Constrained-Attention Multiple-Instance Learners for Whole Slide Image Classification (Short) - S16208 | Folami Alamudun <alamudunft@ornl.gov> |
| 11:00am-11:25am | MatFlow: A System for Knowledgebased Novel Materials Design using Machine Learning (Long) - S16210 | Hasan Jamil <jamil@uidaho.edu> |
| 11:25am-11:45pm | Analysis of Correlation between Cold Weather Meteorological Variables and Electricity Outages (Short) - S16209 | Melissa Allen-Dumas <allenmr@ornl.gov> |
| 11:45am-12:10pm | Pseudonymization at Scale: OLCF's Summit Usage Data Case Study (Long) - BigD794 | Rafael Ferreira da Silva <silvarf@ornl.gov> |
| 12:10pm-12:30pm | A Personalized Federated Learning Framework Using Side Information for Heterogeneous Data Classification (Short) - BigD807 | Yupei Zhang <ypzhaang@nwpu.edu.cn> |
| **12:30pm** | **Closing Remarks** | |

Short : 20 min (15min presentation + 5 min Q&A)
Long: 25 min (20min presentation + 5 min Q&A)

## The 6th IEEE International Workshop on Big Data for Financial News and Data
### Date: Wednesday, Dec. 17, 2022, (Japan Standard Time, GMT+9)

*Workshop Chairs: Quanzhi Li, Xiaozhong Liu, Sameena Shah*

| ID | Time | Title | Presenter/Author |
|---|---|---|---|
| S17201 | | *Long Short-Term Memory Neural Network for Financial Time Series* | Carmina Fjellström |
| S17202 | | *NLP for Responsible Finance: Fine-Tuning Transformer-Based Models for ESG* | Stefan Pasch and Daniel Ehnes |
| S17207 | | *Applying NLP Techniques to Classify Businesses by their International Standard Industrial Classification (ISIC) Code* | Hannah Bechara, Ran Zhang, Shuzhou Yuan, and Slava Jankin |

| | | | |
|---|---|---|---|
| S17209 | 8:30am - 10:15am | Shortcut Learning in Financial Text Mining: Exposing the Overly Optimistic Performance Estimates of Text Classification Models under Distribution Shift | Matthew Caron |
| N239 | | Integrated Data Mapping Engine (DaME) for Financial Services | Shubhi Asthana and Ruchi Mahindru |
| S17205 | | Patentopia: A multi-stage patent extraction platform with disambiguation for certain semantic challenges | Andrea Belz, Alexandra Graddy-Reed, FNU Shweta, Aleksandar Giga, and Shivesh Meenaksh Murali |
| | | Coffee Break | |
| BigD765 | 10:30am - 12:30pm | SFPAD: Sentiment-driven Firm Pricing Anomaly Detection Algorithm and Application | Senyuan Zheng, Vincent CS Lee, and Zhaolin Guan |
| BigD838 | | Relationship between environmental management efforts and corporate performance | Hiroshi Sugeno and Aki-Hiro Sato |
| S17204 | | Hierarchical Multi-task Learning for Enterprise Risk Detection from Financial Documents | Xurui Li, Kaiyuan Liu, Rui Zhu, Yangyang Kang, Changlong Sun, Kaisong Song, and Xiaozhong Liu |
| S17208 | | Bitcoin Evolution Analytics: Twitter Sentiments to Predict Price Change as Bearish or Bullish | Naman Srivastava, Omkar Gowda, Shreyas Bulbule, Siddharth Bhandari, and Animesh Chaturvedi, |
| N203 | | Enhancement to Training of Bidirectional GAN: An Approach to Demystify Tax Fraud | Priya Mehta, Sandeep Kumar, Ravi Kumar, and Sobhan Babu Ch |
| BigD719 | | Dynamic Stop-Loss Approach for Short Term Trades using Deep Learning | Iromie Samarasekara, Oshan Mendis, Sapumal Ahangama, and Ajantha Atukorale |
| | | Closing Remarks | |

| IEEE Big Data Workshop on Multimodal Big Data<br>Dec. 17 8:00am-12:30pm<br>(MMBD 2022, Online)<br>*Workshop Chair: Duoduo Liao*<br>*Web link: https://www.comstar-tech.org/workshops/workshop_2022_MMBD.html* | | |
|---|---|---|
| **Time (JST)** | **Title** | **Presenter/Author** |
| 8:00-8:15 | On the Use of Convolutional Neural Networks for Palm Vein Recognition (S18209) | Mohamed Eladlani, Larbi Boubchir, Said Si Kaddour, and Boubaker Daachi |
| 8:15-8:40 | Radiopaths: Deep Multimodal Analysis on Chest Radiographs (S18204) | Mohammad Kohankhaki, Ahmad Ayad, Mahdi Barhoush, and Anke Schmeink |
| 8:40-9:20 | **Keynote**<br>Social Event Mining and Impact Forecasting with Heterogeneous Spatiotemporal Data | **Dr. Kaiqun Fu** |
| 9:20-9:25 | **Cyber Coffee Break** | |

| 9:25-9:50 | Twitter Bot Identification: An Anomaly Detection Approach (S18205) | Lulwah Alkulaib, Yanshen Sun, Lei Zhang, and Chang-Tien Lu |
|---|---|---|
| 9:50-10:15 | Multimodal Lyrics-Rhythm Matching (S18208) | Callie Liao, Duoduo Liao, and Jesse Guessford |
| 10:15-10:40 | Simulating Fake News Dissemination on Twitter with Multivariate Hawkes Processes (S18206) | Yichen Jiang and Michael Porter |
| 10:40-11:05 | EBEM: An Enhanced Bi-encoder Model for Word Sense Disambiguation (S18202) | Xianglong Xiao, Hongguang Xu, and Ke Xu |
| 11:05-11:10 | **Cyber Coffee Break** | |
| 11:10-11:25 | Multimodal Deep Homography Estimation Using a Domain Adaptation Generative Adversarial Network (S18203) | Thomas Pouplin, Hughes Perreault, Benoit Debaque, Marc-Antoine Drouin, Nicolas Duclos-Hindie, and Simon Roy |
| 11:25-11:40 | Truth or Fiction: Multimodal Learning Applied to Earnings Calls (S18211) | Jamshed Kaikaus, Jessen Hobson, and Robert Brunner |
| 11:40-11:55 | Introduction to Multi-Modal Stylometry for Authorship Identification: Estimating the identity of Satoshi Nakamoto (S18212) | Glory Adebayo and Roman Yampolskiy |
| 11:55-12:10 | Attention-aware multi-modal RNN for deception detection (S18210) | Shun-Wen Hsiao and Cheng-Yuan Sun |
| 12:10-12:25 | SQ2SV: Sequential Queries to Sequential Videos retrieval (S18207) | Injin Paek, Nayoung Choi, Seongjin Ha, Yuheun Kim, and Min Song |
| 12:25-12:30 | **Closing Remarks** | |

<table>
<tr><td colspan="3" style="background:green">

## 2<sup>nd</sup> Workshop on Knowledge Graphs and Big Data
## In Conjunction with IEEE Big Data 2022
## Saturday, Dec. 17, 2022
*Workshop Co-Chairs: Yuan An, Jane Greenberg, Yuan Ling, Alex Kalinowski*
</td></tr>
</table>

| Time (Japan Standard Time) | Title | Presenter/Author |
|---|---|---|
| 9:00-9:05am | Welcome | Yuan An |
| 9:05-9:25am | Learning Framework of Entity and Entity Type Composition Representation for Knowledge Graph Completion (Paper Id: S19201) | Runyu Ni, Hiroki Shibata, and Yasufumi Takama |
| 9:25-9:45am | Workflow for Domain- and Task-Sensitive Curation of Knowledge Graphs, with Use Case of DRKG (Paper Id: S19202) | Kara Schatz, Daniel Korn, Alexander Tropsha, and Rada Chirkova |
| 9:45-10:05am | Development and Evaluation of Embedding Methods for Graphs with Multi Attributes (Paper Id: S19205) | Miyu Fujii, David Taingngin, Keiichiro Yamamura, Nozomi Hata, Hiroki Kai, Ryuji Noda, Hiroki Ishikura, Tatsuru Higurashi, and Katsuki Fujisawa |
| 10:05-10:25am | A Stochastic Geo-spatiotemporal Bipartite Network to Optimize GCOOS Sensor Placement Strategies (Paper Id: S19206) | Ted Holmberg, Elias Ioup, and Mahdi Abdelguerfi |
| 10:25-10:45am | EntityRank: Unsupervised Mining of Bilingual Named Entity Pairs from Parallel Corpora for Neural Machine Translation (Paper Id: S19207) | Min Zhang, Song Peng, Hao Yang, Yanqing Zhao, XiaoSong Qiao, Junhao Zhu, Shimin Tao, Ying Qin, and Yanfei Jiang |
| 10:45-10:55am | **Coffee Break** | |
| 10:55-11:15am | Wikidata-lite for Knowledge Extraction and Exploration (Paper Id: S19208) | Phuc Nguyen and Hideaki Takeda |
| 11:15-11:35am | Exploring Pre-Trained Language Models to Build Knowledge Graph for Metal-Organic Frameworks (MOFs) (Paper Id: S19204) | Alex Kalinowski |
| 11:35-11:55am | Optimizing the Descendant-Aware Clustering Parameters (Paper Id: BigD225) | Sukhwan Jung and Aviv Segev |
| 11:55-12:15pm | Ripl: Document-Level Event Argument Extraction via Role-aware Interactive Pointer Labeling Network (Paper Id: BigD368) | Li Yuanzhuo |

| 12:15-12:30pm | Closing Remarks and Future Suggestions |
|---|---|

<br>

## The Sixth IEEE International Workshop on Benchmarking, Performance Tuning and Optimization for Big Data Applications (BPOD), Saturday, December 17

*Workshop Chairs: Zhiyuan Chen, Jianwu Wang, Feng Chen, Junqi Yin*

Workshop website: https://userpages.umbc.edu/~jianwu/BPOD

| Time (US Eastern Standard Time) | Title | Presenter/Author |
|---|---|---|
| Session 1 (Analytics), Session Chair: Zhiyuan Chen |||
| 2:00-2:05 pm | **Opening Remark** | Workshop Chairs |
| 2:05-2:20 pm | S20208: Sparse Spatio-Temporal Neural Network for Large-Scale Forecasting | Eamon Bracht, Volodymyr Kindratenko, and Robert Brunner |
| 2:20 pm -2:40 pm | S20207: FaST : An Efficient Framework For Visualizing Large-Scale Time Series | Manele Ait Habouche, Mickael Kerboeuf, Goulven Guillou, and Jean-Philippe Babau |
| 2:40 pm-2:55 pm | S20202: Distributed Node Classification with Graph Attention Networks | Roberto Corizzo and Terry Slenn |
| 2:55 pm-3:15 pm | BigD576: Low-Energy Convolutional Neural Networks (CNNs) using Hadamard Method | Varun Mannam |
| Break |||
| Session 2 (Analytics, Workflow & Benchmark), Session Chair: Jianwu Wang |||
| 3:20-3:40 pm | S20205: CH3: A Mixed Workload Benchmark for Scalable NoSQL | Mehnaz Tabassum Mahin, Bo-Chun Wang, Kamini Jagtiani, Michael Carey, and Keshav Murthy |
| 3:40-4 pm | S20210: Discovering Limitations of Image Quality Assessments with Noised Deep Learning Image Sets | Wei Dai and Daniel Berleant |
| 4-4:15 pm | S20201: Leveraging Reinforcement Learning for Task Resource Allocation in Scientific Workflows | Jonathan Bader, Nicolas Zunker, Sören Becker, and Odej Kao |
| 4:15-4:35 pm | BigD364: TAED: Topic-Aware Event Detection | Yan Liang and Christan Grant |
| Break |||
| Session 3 (Optimization and Analytics), Session Chair: Feng Chen |||
| 4:40 pm-5 pm | S20209: Outlier Detection Based on Stacked Autoencoder and Gaussian Mixture Model | Jing Li, Pengbo Lv, Huijun Li, and Wanghu Chen |
| 5-5:20 pm | S20203: ReAssigner: A Plug-and-Play Virtual Machine Scheduling Intensifier for Heterogeneous Requests | Haochuan Cui, Junjie Sheng, Bo Jin, Yiqiu Hu, Li Su, Lei Zhu, Wenli Zhou, and Xiangfeng Wang |
| 5:20-5:40 pm | BigD682: Text Spotting in Natural Scenes Based on Feature Pyramid Neural Network | Chien-Kang Huang and Guojhen Liou |
| 5:40-6 pm | BigD399: Usage-Based Decision Tree for Telco Churn Prediction | Bing NI and Brandon WU |
| Closing Remarks |||

<br>

## Applications of Big Data Technology in the Transport Industry

*Workshop Chair: John Easton, j.m.easton@bham.ac.uk*

| Time | Title | Presenter/Author |
|---|---|---|
| 14:05 – 14:30 | BigD345 – IoT for Rail Transportation: The Case of Railigent | Simon Frederic Dietlmeier, Rob Jago Floetgen, Julius Bock, and Florian Urmetzer |
| 14:30 – 14:55 | S21201 – Uncertainty of Traffic Congestion Estimation Using Nationwide Pseudo Trip Data and Agent-Based Simulation | Aayush Tewari, Yanbo Pang, and Yoshihide Sekimoto |
| 14:55 – 15:20 | S21203 – Towards STUB 2.0: Using Graph-Based World States in Hyperledger Besu to Facilitate Distributed Transport Ticketing | Joseph Preece, Christopher Morris, and John Easton |
| 15:20 – 15:45 | S21202 – Analysis of Airfare during Pandemic: A Multi-Agent Based Modeling Approach | Abdul Mutakabbir, Chung-Horng Lung, and Samuel A. Ajila |
| 16:00 – 16:20 | Coffee Break ||
| 16:20 – 16:45 | S21205 – Digitalization and the Transport Industry – Are we Ready? | John Easton |

| 16:45 – 17:10 | BigD347 – Analysis of route efficiency in city bus transportation vulnerable areas, Focusing on Yeongjongdo, Incheon Metropolitan City | Yoon Nam Sik, Choi Inn Jun, and Kang Sung Woo |
|---|---|---|
| 17:10 – 17:35 | S21204 – A Graph-based Analysis Approach to Cluster Lifetime Dynamics | Ivens Portugal, Paulo Alencar, and Donald Cowan |
| 17:35 – 18:00 | BigD436 – Developing Origin and Destination Matrices in Megapolitan Areas Leveraging Big Data: A Case Study of the Greater Jakarta Metropolitan Area | Herawati Basirpuly, Idwan Santoso, Yugi M.Hartiman, Atik Siti Kuswati, Daniella Daniella, Andrean Maulana, Russ Bona Frazila, Irawati Adriani, and Rita Pasaribu |
| 18:00 | **Closing Remarks** | |

## 3rd International Workshop on User Understanding from Big Data
*WorkshopChairs: Claire Ding; Wutao Wei*

| Time | Title | Presenter/Author |
|---|---|---|
| 8:35-8:55 | ***"Prospect-Net: Top-K Retrieval Problem Using Prospect Theory"*** | Reza Yousefi Maragheh |
| 9:00-9:20 | ***"Understanding Levels of Automation in Human-Machine Collaboration"*** | Glaucia Melo |
| 9:25-9:45 | ***"Posteriori Trust: Behavior driven trust on a network graph at scale"*** | Prince Paulraj, Elijah Hall, Ana Armenta, and Mark Austin |
| 9:50-10:10 | ***"A Short Survey on the User Cold Start Problem in Recommender Systems: Metadata and Meta-Learning Methods"*** | Hao Jiang |
| 10:15-10:30 | **Coffee Break** | |
| 10:30-10:50 | ***"Blockchain based Secure Group Data Collaboration in Cloud with Differentially Private Synthetic Data and Trusted Execution Environment"*** | Uzair Javaid |
| 10:55-11:15 | ***"Where did you tweet from? Inferring the origin locations of tweets based on contextual information"*** | Rabindra Lamsal |
| 11:20-11:40 | ***"KGIQ: Scalable Translation of User-Specified Examples into Knowledge-Graph Queries"*** | Jing Ao |
| 11:45-12:05 | ***"Preserving Privacy Mining of Big Temporal Co-occurrence Patterns"*** | Anifat Olawoyin |
| 12:10-12:15 | **Closing Remarks** | |

## The Sixth IEEE Workshop on Human-in-the-loop Methods and Future of Work in BigData (HMData 2022)
*Workshop Chair: Senjuti Basu Roy (New Jersey Institute of Technology), Alex Quinn (Purdue University), Atsuyuki Morishima (University of Tsukuba)*
*The latest program is available at https://humanmachinedata.org*

| Time | Title | Author(s) |
|---|---|---|
| 8:50 | **Opening & Announcements** | |
| 9:00 | **Keynote:** Ming Yin (Purdue University) | |
| 9:50 | **Short Break** | |
| 10:00 | [S24206] An Adaptive Human-in-the-loop Approach to Emission Detection of Additive Manufacturing Processes and Active Learning with Computer Vision | Xiao Liu, Alan F. Smeaton, and Alessandra Mileo |
| 10:20 | [S24212] Image Geolocation by Non-Expert Crowd Workers with an Expert Strategy | Seungun Kim, Masaki Matsubara, and Atsuyuki Morishima |
| 10:32 | [S24214] A SHAP-based Active Learning Approach for Creating High-Quality Training Data | Nailcan Kara, Yagiz Levent Gume, Umit Tigrak, Gokce Ezeroglu, Serdar Mola, Omer Burak Akgun, and Arzucan Ozgur |
| 10:44 | [S24211] Graph Neural Network Tells Us Who is the Communication Enhancer | Hiroaki Tanaka, Yuuma Jitsunari, Wataru Yamada, Keiichi Ochiai, Shoko Wakamiya, and Eiji Aramaki |
| 10:56 | **Break** | |

| Time | Title | Presenter/Author |
|---|---|---|
| 11:20 | [S24204] Toward Data Cleaning with a Target Accuracy: A Case Study for Value Normalization | Adel Ardalan, Derek Paulsen, Amanpreet Singh Saini, Walter Cai, and AnHai Doan |
| 11:40 | [S24210] Efficient Evaluation of AI Workers for the Human+AI Crowd Task Assignment | Tomoya Kanda, Hiroyoshi Ito, and Atsuyuki Morishima |
| 12:00 | [S24215] Multi-Armed Bandit Approach to Qualification Task Assignment across Multi Crowdsourcing Platforms | Yunyi Xiao, Yu Yamashita, Hiroyoshi Ito, Masaki Matsubara, and Atsuyuki Morishima |
| 12:12 | [S24219] DelphAI: A human-centered approach to time-series forecasting | Kristina L. Kupferschmidt, Joshua G. Skorburg, and Graham W. Taylor |
| 12:24 | [S24209] Crowd-Aided Anomaly Detection in Surveillance Videos | Ryuya Itano, Tomoya Nohara, and Takahiro Koita |
| 12:36 | **Lunch Break** | |
| 14:00 | [S24205] Explainable Recommendation Using Knowledge Graphs and Random Walks | Kaname Muto, Satoshi Oyama, and Itsuki Noda |
| 14:20 | [S24207] Owro: A Novel Robot for Sitting Posture Training Based on Adaptive Human Robot Interaction | Katia Bourahmoune, Karlos Ishac, Marc Carmichael, and Toshiyuki Amagasa |
| 14:40 | [S24221] Improving Character Recognition by the Crowd Workers via Corrective Feedback | Bijoly Saha Bhattacharya, Biswajit Mandal, Arindam Biswas, and Malay Bhattacharyya |
| 14:52 | [S24216] PostMe: Unsupervised Dynamic Microtask Posting For Efficient and Reliable Crowdsourcing | Ryo Yanagisawa, Susumu Saito, Teppei Nakano, Tetsunori Kobayashi, and Tetsuji Ogawa |
| 15:04 | [S24208] Multi Speaker Identification System using Crowdsourcing-based Method | Kazumu Nakahira, Shun Kuroiwa, and Takahiro Koita |
| 15:16 | **Closing Remarks** | |

## Big Food, Nutrition and Environment Data Management and Analysis
*WorkshopChairs: Tome Eftimov, Giulia Menichetti*

| Time | Title | Presenter/Author |
|---|---|---|
| 14:00-14:25 | Invited talk I | Babak Ravandi |
| 14:25-14:50 | Invited talk II | Chris Evelo |
| 14:50-15:05 | S25204 - Using Big Data and Machine Learning for Multi-layered Surveillance for Healthy Food Environment and Diet | Fares Belkhiria |
| 15:05-15:20 | S25203 - Heterogeneity in feature importance and prediction performance for sales at the market and store levels: the case of branded yogurt products in Quebec | Cameron McRae |
| 15:20-15:35 | S25206 - The Ontological How and Why - Action and Objective - of Planned Processes in the Food Domain | Tarini Naravane |
| 15:35-15:50 | S25208 - Learning Robust Food Ontology Alignment | Riste Stojanov |
| 15:50-16:05 | **Coffee Break** | |
| 16:05-16:30 | Invited talk III | Aleksandra Kovachev |
| 16:30-16:55 | Invited talk IV | Joan Capdevila Pujol |
| 16:55-17:10 | S25207 - SciFoodNER: Food Named Entity Recognition for Scientific Text | Gjorgjina Cenikj |
| 17:10-17:25 | S25205 - Predefined domain specific embeddings of food concepts and recipes: A case study on heterogeneous recipe datasets | Gordana Ispirova |
| 17:25-17:40 | S25201 - Privacy-preserving & machine-learned catchment models for national dietary surveillance via digital footprint data | Gavin Long |
| 17:40-17:55 | S25202 - Ill-fated interactions: modeling complaints on a food waste fighting platform | Georgiana Nica-Avram |
| **17:55-18:00** | **Closing Remarks** | |

## The Fifth Workshop on Big Data for Economic and Business Forecasting
*WorkshopChairs: Wei Shang, Jue Wang*

| Time | Title | Presenter/Author |
|---|---|---|

| Time | Title | Presenter/Author |
|---|---|---|
| 14:00-14:25 | BigD534 - Restaurant Sales Forecasting with Feature Interaction-learning Mechanism-based Neural Network Model | Xin Yang |
| 14:25-14:50 | S26202 - Using CNN-LSTM Model for Weather Forecasting | Samuel A. Ajila |
| 14:50-15:15 | BigD503 - Aggregate Learning for Mixed Frequency Data | Daisuke Moriwaki |
| 15:15-15:40 | S26203 - DVM-CAR: A Large-Scale Automotive Dataset for Visual Marketing Research and Applications | Jingmin Huang |
| 15:40-16:05 | BigD349 - Forecast of China's electricity consumption Considering the fast growth of New Infrastructure | Tianfeng Li |
| 16:05-16:20 | **Coffee Break** | |
| 16:20-16:45 | S26201 - A Semisupervised Approach to Predicting a Twitch Streamer's Growth based on the Game Streamed | Samuel A. Ajila |
| 16:45-17:10 | BigD637 - Comparison of Consensus Building Processes by Different Reach Patterns on Opinion Dynamics | Makoto Fujii |
| 17:10-17:35 | BigD640 - Influence of a leader and in-group pressure in opinion dynamics model | Yunosuke Hori |
| 17:35-18:00 | BigD266 - Bayesian Ensembled Knowledge Extraction Strategy for Online Portfolio Selection | Aviv Segev |
| | **Closing Remarks** | |

| The 5th International Workshop on Big Media Dataset Construction, Management and Applications | | |
|---|---|---|
| *WorkshopChairs: Chaoqun Hong, Xiaoqin Zhang* | | |
| **Time** | **Title** | **Presenter/Author** |
| 13:00 – 13:20 | #S27201, Scene Text Recognition with Heuristic Local Attention | Tianlong Ma |
| 13:20 – 13:40 | #S27202, Dual Priors Network for RGB-D Salient Object Detection | Yuewang Xu |
| 13:40 – 14:00 | #S27203, Self-Supervised Person Re-Identification with Channel-Wise Transformer | Zian Ye |
| 14:00 – 14:20 | #S27204, Personalized Federated Contrastive Learning | Yupei Zhang |
| 14:20 – 14:40 | **Coffee Break** | |
| 14:40 – 15:00 | #S27205, EFD:A New Benchmark for Load Identification and Energy Disaggregation | Ziheng Sun |
| 15:00 – 15:20 | #S27206, MS2A: A Multi-Scale Spatial Aggregation Framework for Visual Place Recognition | Ge Cui |
| 15:20 – 15:40 | #S27207, ADIU: An Antiquarian Document Image Unwarping Dataset | Shenlu Liu |
| 15:40 – 16:00 | #S27209, Deformable Attention improved Human Pose Estimation with Monocular | Yuhong Xie |
| 16:00 – 16:20 | **Closing Remarks** | |

**Please note this is for the three workshops: BDA4CID, BDA4HM and BDA4S. Together these make a two-day workshop of 39 papers. Day 1 (17th December) will be in-person presentations in Osaka. Day 2 (18th December) will be virtual presentations. All times are on local time Osaka (GMT/UTC+9).**

# Day 1: Physical – ideally 17th December 2022

| 6th International Workshop on Big Data Analytics for Cyber Intelligence and Defense (BDA4CID) | | |
|---|---|---|
| *WorkshopChairs: Stephen McGough, Amir Atapour-Abarghouei, David Wall* | | |
| **Time** | **Title** | **Presenter/Author** |
| 08:00 – 08:20 | Introduction to Physical workshop | Stephen McGough |
| 08:20 – 08:45 | A User and Entity Behavior Analytics Log Data Set for Anomaly Detection in Cloud Computing (S29201) | **Max Landauer**, Florian Skopik, Georg Hoeld, and Markus Wurzenberger |
| 08:45 – 09:10 | Building Cybersecurity Ontology for Understanding and Reasoning Adversary Tactics and Techniques (S29207) | **Chiao-Cheng Huang**, Pei-Yu Huang, Ying-Ren Kuo, Guo-Wei Wong, Yi-Ting Huang, Yeali S. Sun, and Meng Chang Chen |
| 09:10 – 09:35 | Simple and Efficient Identification of Personally Identifiable Information on Public Website (S29211) | **Caitlin Brown** and Charles Morisset |

| 09:35 – 10:00 | Factor Detection Task of Cyberbullying Using the Deep Learning Model (BigD809) | **Yu-Hsuan Wu**, Sheng-Wei Huang, Wei-Yi Chung, Chen-Chia Yu, and Jheng-Long Wu |
|---|---|---|
| 10:00 – 10:30 | **Coffee Break** | |

## 1st International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2022)
*WorkshopChairs: Stephen McGough, Matthew Forshaw, Amir Atapour-Abarghouei*

| Time | Title | Presenter/Author |
|---|---|---|
| 10:30 – 10:55 | Comparing Association Rules and Deep Neural Networks for Heart Disease Prediction (S30207) | **Carlos Ordonez**, Ian Fund, and Ladjel Bellatreche |
| 10:55 – 11:20 | Explainable Deep Learning to Profile Mitochondrial Disease Using High Dimensional Protein Expression Data (S30211) | **Atif Khan**, Conor Lawless, Amy Vincent, Satish Pilla, Sushanth Ramesh, and Stephen McGough |
| 11:20 – 11:45 | Discovering Similar Spike Patterns in High Dimensional Biomedical Signals (S30208) | Sikder Tahsin Al-Amin, Robin Varghese, David Lloyd, Maria A. Gonzalez-Gonzalez, Mario I. Romero-Ortega, and **Carlos Ordonez** |
| 11:45 – 12:10 | Monitoring and Improving Personalized Sleep Quality from Long-Term Lifelogs (S30213) | **Wenbin Gan**, Minh-Son Dao, and Koji Zettsu |
| 12:10 – 12:35 | Tracking trajectories of multiple long-term conditions using dynamic patient-cluster associations (S30210) | Ron Kremer, Syed Raza, Fabiola Eto, John Casement, Christian Atallah, Sarah Finer, Dennis Lendrem, Michael Barnes, Nick Reynolds, and **Paolo Missier** |
| 12:35 – 14:00 | **Lunch Break** | |
| 14:00 – 14:25 | Individual based epidemiological model of COVID19 using location data (BigD342) | **Yoriyuki Yamagata**, Shunki Takami, Keisuke Yamazaki, Tomoki Nakaya, and Masaki Onishi |
| 14:25 – 14:50 | Siamese Neural Networks for Skin Cancer Classification and New Class Detection using Clinical and Dermoscopic Image Datasets (BigD517) | Michael Luke Battle, Amir Atapour-Abarghouei, and **Stephen McGough** |
| 14:50 – 15:15 | The Impact of COVID-19 on Human Mobility: A Case Study on New York (BigD548) | Xinchen Hao, **Renhe Jiang**, Jiewen Deng, and Xuan Song |

## 2nd International Workshop on Big Data Analytics for Sustainability (BDA4S)
*WorkshopChairs: Stephen McGough, Gavin Shaddick, Hao Dong*

| Time | Title | Presenter/Author |
|---|---|---|
| 15:15 – 15:40 | Optimizing a domestic battery and solar photovoltaic system with deep reinforcement learning (S31202) | Alexander Kell, **Stephen McGough**, and Matthew Forshaw |
| 15:40 – 16:05 | Analyzing the Effects of Topics Underlying Companies' Financial Disclosures about Risk Factors on Prediction of ESG Risk Ratings: Emphasis on BERTopic (S31205) | Min Gyeong Kim, Kyu Sung Kim, and **Kun Chang Lee** |
| 16:05 – 16:30 | **Coffee Break** | |
| 16:30 – 16:55 | AI Approaches for Electricity Price Forecasting Stable/Unstable Markets: EU Improvement Project (BigD734) | Florian Chauvet, **Ladjel Bellatreche**, and Carlos Santos-Silva |
| 16:55 – 17:20 | A Data Integration Approach to Estimating Personal Exposures to Air Pollution (S31206) | Matthew Thomas, **Gavin Shaddick**, David Topping, Karyn Morrissey, Thomas Bannan, Mike Diessner, Ruth Bowyer, Stefan Siegert, Hugh Coe, James Evans, Fernando Benitez-Paez, and James Zidek |
| 17:20 – 17:45 | Using Contextual Text Mining and Ontology Methods to Establish a Novel Technology Trend and Associative Analysis Framework for Sustainable Energy Development in Taiwan (S31210) | **Yi-Hao Hsiao**, Chia-Yi Chuang, Megn-Chi Huang, Chia-Lee Yang, and Jyh-Horng Wu |
| 17:45 – 18:10 | AI applications in forest monitoring need remote sensing benchmark datasets (S31212) | **Emily Lines**, Matt Allen, Carlos Cabo, Kim Calders, Amandine Debus, Stuart Grieve, Milto Miltiadou, Adam Noach, Harry Owen, and Stefano Puliti |

| Time | Title | Presenter/Author |
|---|---|---|
| 18:10 – 18:35 | Residual value prediction using deep learning for a circular economy (BigD535) | **Edvin Listo Zec**, Olof Mogren, Ann-Charlotte Mellquist, Sarah Fallahi, and Peter Algurén |
| 18:35 – 19:00 | The Impact of Multicore CPUs on Eco-Friendly Query Processors in Big Data Warehouses (BigD732) | **Ayoub Bouhatous**, Ladjel Bellatreche, El Hassan Abdelwahed, and Carlos Ordonez |
| 19:00 | **Closing Remarks** | |

# Day 2: Virtual – ideally 18<sup>th</sup> December 2022

**6th International Workshop on Big Data Analytics for Cyber Intelligence and Defense (BDA4CID)**
*WorkshopChairs: Stephen McGough, Amir Atapour-Abarghouei, David Wall*
*And*
**1st International Workshop on Big Data Analytics for Health and Medicine (BDA4HM 2022)**
*WorkshopChairs: Stephen McGough, Matthew Forshaw, Amir Atapour-Abarghouei*
*And*
**2<sup>nd</sup> International Workshop on Big Data Analytics for Sustainability (BDA4S)**
*WorkshopChairs: Stephen McGough, Gavin Shaddick, Hao Dong*

| Time | Title | Presenter/Author |
|---|---|---|
| 08:00 – 08:20 | Welcome to the day | Stephen McGough |
| 08:20 – 08:45 | Natural Language Interface for Process Mining Queries in Healthcare (BDA4HM - S30212) | **Hangu Yeo**, Elahe Khorasani, Vadim Sheinin, Irene Manotas, Ngoc Phuoc An Vo, Octavian Popescu, and Petros Zerfos |
| 08:45 – 09:10 | AsdClaims: Twitter Dataset of Claims on Autism Spectrum Disorder (BDA4HM - S30209) | **Michelle Kim**, Junghwan Kim, Bryan Kim, Kristen Jonhson, and Jee-In Kim |
| 09:10 – 09:35 | Application of Time Series Analysis Methods to Assess the impact of COVID-19 on Carbon Dioxide Emission Reduction in Connecticut (BDA4S - S31203) | Katherine Kiernan, Hao Wu, and **Yulei Pang** |
| 09:35 – 10:00 | RaCAViz: Interactive Visualizations for Rapid Carbon Assessment (BDA4S - S31207) | **Du Nguyen** and Van Vung Pham |
| 10:00 – 10:30 | **Coffee Break** | |
| 10:30 – 10:55 | Fraud Detection Using Optimized Machine Learning Tools Under Imbalance Classes (BDA4CID - S29209) | **Mary Isangediok** and Kelum Gajamannage |
| 10:55 – 11:20 | PMFL: Partial Meta-Federated Learning for heterogeneous tasks and its applications on real-world medical records (BDA4HM - BigD341) | **Tianyi Zhang,** Shirui Zhang, Ziwei Chen, Yoshua Bengio, and Dianbo Liu |
| 11:20 – 11:45 | MATE: Summarizing Alerts to Interpretable Outcomes with MITRE ATT&CK (BDA4CID - S29203) | **Derek Lin** |
| 11:45 – 12:10 | Cross-Lingual Propaganda Detection (BDA4CID - S29202) | **Wenshan Zhang** and Xi Zhang |
| 12:10 – 12:35 | Federated Learning Attacks and Defenses: A Survey (BDA4CID - S29206) | **Yao Chen**, Yijie Gui, Hong Lin, Wensheng Gan, and Yongdong Wu |
| 12:35 – 14:00 | **Lunch Break** | |
| 14:00 – 14:25 | Fast Illegal Webpage Detection Algorithm Based on Massive Domain Name Resolution Records (BDA4CID - BigD206) | Yue Su, **Botao Peng**, and Xiaodong Li |
| 14:25 – 14:50 | Pavement Condition Detection Method Based on Time-Frequency Features and Capsule Neural Network (BDA4S - S31211) | Wanghu Chen, **Pengbo Lv**, and Jing Li |
| 14:50 – 15:15 | Analysis of Label-Flip Poisoning Attack on Machine Learning Based Malware Detector (BDA4CID - S29205) | **Kshitiz Aryal**, Maanak Gupta, and Mahmoud Abdelsalam |
| 15:15 – 15:40 | Blood Cells Counting and Localisation through Deep Learning Object Detection (BDA4HM - S30201) | **Francesco Mercaldo**, Fabio Martinelli, Antonella Santone, and Mario Cesarelli |
| 15:40 – 16:05 | Exploiting Liver CT scans in Colorectal Carcinoma genomics mutation classification (BDA4HM - S30203) | **Daniele Perlo**, Luca Berton, Alessia Delpiano, Francesca Menchini, Stefano Tibaldi, Marco Grosso, and Paolo Fonio |
| 16:05 – 16:30 | **Coffee Break** | |
| 16:30 – 16:55 | On the Use of Conventional Neural Networks for COVID-19 Detection in CT-Scan Images : A Comparative Study and Performance Analysis (BDA4HM - S30206) | **Menglin Niu**, Larbi Boubchir, and Boubaker Daachi |
| 16:55 – 17:20 | Smart Grid monitoring through Deep Learning for Image-based Automatic Dial Meter Reading (BDA4S - S31201) | **Francesco Mercaldo**, Fabio Martinelli, and Antonella Santone |

| 17:20 – 17:45 | Improving conversational recommender systems via knowledge graph-based semantic fusion with historical interaction data (BDA4CID - S29204) | **Thamizhiniyan Pugazhenthi** and Huizhi Liang |
|---|---|---|
| 17:45 – 18:10 | Threat Modeling for Machine Learning-Based Network Intrusion Detection Systems (BDA4CID - S29210) | **Huda Ali Alatwi** and Charles Morisset |
| 18:10 – 18:35 | A systematic literature review on machine learning for electricity market agent-based models (BDA4S - S31204) | **Alexander Kell**, Stephen McGough, and Matthew Forshaw |
| | **Closing Remarks** | |

<br>

| Second International Workshop on<br>Data science for equality, inclusion and well-being challenges<br>DS4EIW 2022@IEEE Big Data 2022 – ONLINE event<br>December 20, 2022 (9:00a.m – 1:00p.m (CET, GMT+1)<br>Workshop Chairs: Tania Cerquitelli, Genoveva Vargas-Solar, Arianna Montorsi,<br>Claudia De Giorgi<br>Session Chairs: Genoveva Vargas-Solar and Tania Cerquitelli |||
|---|---|---|
| **Time** | **Title** | **Presenter/Author** | **Paper ID** |

| Time | Title | Presenter/Author | Paper ID |
|---|---|---|---|
| 9:00 – 9:05 | **Opening** | | |
| 9:05 -- 9:25 | Peers and Stars: the role of gender among coinventors | Federico Caviggioli, Alessandra Colombelli, and Chiara Ravetti | S34201 |
| 9:25 -- 9:45 | Obtaining Dyadic Fairness by Optimal Transport | Moyi Yang, Junjie Sheng, Wenyan Liu, Bo Jin, Xiaoling Wang, and Xiangfeng Wang, | S34203 |
| 9:45 -- 10:05 | Identifying Imbalance Thresholds in Input Data to Achieve Desired Levels of Algorithmic Fairness | Mariachiara Mecati, Andrea Adrignola, Antonio Vetrò, and Marco Torchiano | S34205 |
| 10:05 -- 10:25 | Better SVM-based Record Linkage by Addressing Hidden Bias | Christina Makri, Alexandros Karakasidis, and Evaggelia Pitoura | S34206 |
| 10:25 -- 10:45 | Psychological Attention-based Analytics of Multivariate Campus Behaviors of University Students | Wanghu Chen, Jilong Yao, Yang Liu, Jing Li, Yan Sun, and Chunyu Pang | S34207 |
| 10:45 -- 11:05 | Physical and mental health of university staff during the Covid-19 pandemic | Alessandra Colombelli, Greta Temporin, Francesco Serraino, and Tania Cerquitelli | S34209 |

| Time | Title | Author | Code |
|---|---|---|---|
| **11:05 – 11:15** | **Coffee Break** | | |
| 11:15 – 11:35 | Trade less Accuracy for Fairness and Trade-off Explanation for GNN | Yazheng Liu, Xi Zhang, and Sihong Xie | S34211 |
| 11:35 – 11:55 | Nondiscriminating Dropout Prediction beyond Algorithmic Fairness: Ensuring Coverage in Preprocessing Pipelines | Chiara Accinelli, Barbara Catania, and Giovanna Guerrini | S34213 |
| 11:55 – 12:10 | Combining Dynamic Mode Decomposition and Difference-in-Differences in an Analysis of At-Risk Youth | Marc Andrew Choi, Siyu Huang, Hengyuan Qi, Marco Scialanga, Emerson McMullen, Axel Sanchez Moreno, Yifei Lou, Andrea L. Bertozzi, and P. Jeffrey Brantingham | S34214 |
| 12:10 – 12:30 | Systematic Discovery of Bias in Data | John Wamburu, Girmaw Abebe Tadesse, Celia Cintas, Adebayo Oshingbesan, Tanya Akumu, and Skyler Speakman | S34204 |
| 12:10 – 12:50 | Promoting equity, diversity and inclusion: policies, strategies and future directions in higher education, research communities and business | Genoveva Vargas-Solar, Tania Cerquitelli, Arianna Montorsi, Stefania Salvai, Maria Teresa Sangineti, Jérôme Darmont, and Cécile Favre | S34210 |
| **12:50** | **Closing Remarks** | | |

**IEEE BigData 2022 Workshop on Knowledge Discovery and Data Mining in IT Operations (BigData-IT 2022)**
*Yu Deng, Meng Jiang, Berthold Reinwald, and ChengXiang Zhai*

| Time | Title | Presenter/Author |
|---|---|---|
| 8:30 am | Opening remark | Workshop co-chairs |
| 8:35 am | Keynote speech | Prof. Jiawei Han |
| 9:15 am | S35201: First CE Matters: On the Importance of Long Term Properties on Memory Failure Prediction | Jasmin Bogatinovski |
| 9:30 am | S35203: AIOps Essential to Unified Resiliency Management in Data Lakehouses | Runyu Jin |
| 9:45 am | S35206: Entity Set Co-Expansion in StackOverflow | Yu Zhang |
| 10:00 am | S35207: Storage Capacity Prediction using Population Analytics | Umesh Deshpande |
| 10:15 am | **Coffee Break** | |
| 10:20 am | S35202: Answer Comments As Reviews: Predicting Acceptance By Measuring Valence On Stack Exchange | William Ledbetter |
| 10:40 am | S35204: Fine-Tuning BERT-based Language Models for Duplicate Trouble Report Retrieval | Nathan Bosch |

| Time | Title | Presenter/Author |
|---|---|---|
| 11:00 am | S35205: Improving Model Performance Using Metric-Guided Data Selection Framework | Paulina Toro Isaza |
| 11:20 am | S35208: ContrastBERT: Supervised Contrastive Learning of BERT-Encoded IT logs for Anomaly Classification | Raziuddin Mahmood |
| 11:40 am | BigD234: To test, or not to test: A proactive approach for deciding complete performance test initiation | Omar Javed |
| 12:00 pm | BigD706: Cognitive Metric Monitoring - Characterizing spatial-temporal behavior for anomaly detection | Kunal Jethuri |
| **12:20 pm** | **Closing Remarks** | |

| Big Data Analytics for Humanitarian Crises | | |
|---|---|---|
| *WorkshopChairs: Thilanka Munasinghe (munast@rpi.edu)* | | |
| **Time** | **Title** | **Presenter/Author** |
| 14:00 – 14:45 | Keynote talk by Dr. Jennifer Wei (NASA Goddard) | Dr. Jennifer Wei |
| 14:50 – 15:15 | **ID: S36203**: *TweetDIS: A Large Twitter Dataset for Natural Disasters Built using Weak Supervision* | Ramya Tekumalla, Juan Banda |
| 15:15 – 15:35 | **ID: S36204**: *Feature Selection from Multivariate Time Series Data: A Case Study of Solar Flare Prediction* | Khaznah Alshammari, Shah Muhammad Hamdi, Soukaina Filali Boubrahimi |
| 15:35 – 16:00 | **ID: BigD494**: *An Empirical Study on Characterizing Natural Disasters in Class Imbalanced Social Media Data using Weak Supervisio* | Ramya Tekumalla, Juan Banda |
| 16:00 – 16:20 | **ID: S36202**: *National press monitoring using Natural Language Processing as an early warning signal for prediction of asylum applications flows in Europe* | Ilias Iliopoulos, Nikolaos Kopalidis, Georgios Stavropoulos, and Dimitrios Tzovaras |
| 16:20 – 16:30 | **ID: S36205**: *PROBER: A System for Real-time Propaganda Behavior Analytics on Social Media and Web Data Streams* | Yasas Senarath, Antonios Anastasopoulos, Tonya Thornton, Hemant Purohit, |
| 16:30 – 16:40 | **ID: S36201**: *Digital Financial Literacy, Risk Aversion, and College Students' Online Security Behavior* | Yueqi Li, Ingrid Fisher |
| 16:40 – 17:00 | **Closing Remarks** | |
| | **Lunch** | |

## Fifth International Workshop on the Internet of Things Data Analytics (IoTDA)
*Location: Online (virtual)*
*Workshop Chairs: Eyhab Al-Masri (University of Washington), Yan Bai (University of Washington), Chi-Hua Chen (Fuzhou University, China), Alireza Souri (Halic University, Turkey)*

| Time | Title | Presenter/Author |
|---|---|---|
| 9:00 am – 9:05 am | Opening Remarks and Welcome | |
| 9:10 am – 9:25 am | (S32202)  Towards Normalizing The Design Phase of Data Preprocessing Pipelines For IoT Data | Amal Tawakuli and Thomas Engel |
| 9:30 am – 9:45 am | (S32204)  A Business Workflow Architecture for Predictive Maintenance using Real-Time Anomaly Prediction On Streaming IoT Data | Emrullah Gultekin and Mehmet S. Aktas |
| 9:50 am – 10:05 am | (S32205)  Probabilistic Time Series Forecasting for Adaptive Monitoring in Edge Computing Environments | Dominik Scheinert, Babak Sistani Zadeh Aghdam, Soeren Becker, Odej Kao, and Lauritz Thamsen |
| 10:10 am – 10:25 am | (S27207)  Lossy Predictive Models for Accurate Classification Algorithms | Aekyeung Moon, Seung Woo Son, Hyson Kim, and Minjun Kim |
| **10:30 am – 10:35 pm** | **Closing Remarks** | |

## Ninth International Workshop on High Performance Big Graph Data Management, Analysis, and Mining (BigGraphs 2022)
*Workshop Chairs: Mohammad Al Hasan, Nesreen Ahmed, Shaikh Arifuzzaman, Kamesh Madduri*

| Time | Title | Presenter/Author |
|---|---|---|
| 8:30 am | Opening Remarks | Mohammad Al Hasan |
| 8:35 am | S33201: Data Fusion and Graph Analysis in Fraud Transaction Detection: walkthrough of a case study | Valerio Bellandi and Stefano Siccardi |
| 9:00 am | BigD269: Pipelining Graph Construction and Agent-based Computation over Distributed Memory | Yan Hong and Munehiro Fukuda |
| 9:50 am | BigD387: Functional Ball Dropping: A superfast hypergraph generation scheme | Lilith Hafner, Chase Holdener, and Nicole Eikmeier |
| 10:15 am | S33204: CrunchQA - A Synthetic Dataset for Question Answering over Crunchbase Knowledge Graph | Lifan Yu, Nadya Abdel Madjid, and Djellel Difallah |
| 10:15 – 10:30 am | **Coffee Break** | |
| 10:30 – 11:20 am | BigGraphs 2022 Keynote talk | |
| 11:25 am | BigD543: Node Context Selection in Transformer-Based Graph Representation Learning Models | Tianze Wang, Amir H. Payberah, and Vladimir Vlassov |
| 11:50 am | S33202: Scalable Forecasting in Sensor Networks with Graph Convolutional LSTM Models | Massimiliano Altieri, Roberto Corizzo, and Michelangelo Ceci |
| 12:15 pm | **Closing Remarks** | |

## Digital twins for accelerated discovery of climate and sustainability solutions
*Workshop Chairs: Levente Klein, Mitchaiki Tasubori, Conrad Albrecht*

| Time | Title | Presenter/Author |
|---|---|---|
| 8:30-8.45 am | S37209, *"Identifying Extreme Regimes in Climate-Scale Digital Twins: a Roadmap"* | Eloisa Bentivegna |
| 8.45-9 am | S37201, *"Leak localization in Water Distribution Networks By Directly Fitting the Learning Parameters of a Gaussian Naive Bayes Classifier"* | Ganjour Mazaev |

| Time | Title | Presenter/Author |
|---|---|---|
| 9-9.15 am | S37211, *"Source Localization and Bayesian leak magnitude inference of sparse wireless sensor data to detect fugitive methane leak"* | Levente Klein |
| 9.15-9.30 am | S37205,*"Optimal Sensor Placement for Atmospheric Inverse Modelling"* | Xinchao Liu/Kyongmin Yeo |
| 9:30-9.45 am | S37202, *"Multi-task Learning for Source Attribution and Field Reconstruction for Methane Monitoring"* | Arka Daw/Kyonming Yeo |
| 9.45-10 am | S37208,*"Flood Event Detection from Sentinel 1 and Sentinel 2 Data: Does Land Use Matter for Performance of U-Net based Flood Segmenters?"* | Michal Muszynski |
| 10-10.30 am | **Coffee Break** | |
| 10.30-11.15 am | *Keynote presentation: Digital Twins in IBM* | **Michiaki Tatsubori, IBM** |
| 11.15-11.30 am | S37204, *"Peaks Fusion assisted Early-stopping Strategy for Overhead Imagery Segmentation with Noisy Labels"* | Chenying Liu |
| 11.30-11.45 am | S37203,*"Deep Semantic Model Fusion for Ancient Agricultural Terrace Detection"* | Yi Wang |
| 11.45 am-12 pm | BigD280, "A Multi-objective Residual TrajGRU Model for Wind Field Forecasting" | Wei Zhang |
| 12 pm-12.15 pm | S37206, *"Spatiotemporal River Discharge Interpolation in Ungauged Nodes via Deep Kernel Learning using Representative Topographic Features: A Development Study."* | Refaldi Putra |
| 12.15-12.30 pm | S37210, *"NetZeroCO2, an AI framework for accelerated nature-based carbon sequestration"* | Ademire Ferreira de Silva |
| 12.30-12.35 pm | **Closing Remarks** | |
| | | |

| The 7th IEEE International Workshop on Big Spatial Data | | |
|---|---|---|
| *Workshop Chairs: Farnoush Banaei-Kashani, Chengyang Zhang, Abdeltawab Hendawi, Ashwin Shashidharan* | | |
| **Time (JST)** | **Title** | **Presenter/Author** |
| 8:30am – 8:35am | **Opening Remarks** | |
| 8:35am – 9:35am | Keynote by Dr. Sakai Masato | Dr. Sakai Masato |
| 9:35am – 10:05am | BigD459 - FAGR: Fairness-aware Group Recommendation in Event-based Social Networks | Yuan Liang <liangyuan120@buaa.edu.cn> |
| 10:05am – 10:30am | **Coffee Break** | |
| 10:30am – 11:00am | S40201 - Retrieving Top-N Weighted Spatial k-cliques | Ryosuke Taniguchi <taniguchi.ryosuke@ist.osaka-u.ac.jp> |
| 11:00am – 11:30am | BigD620 - DACMA: Designing space ordering optimizations to scalably manage aerial images | Chamin Nalinda Lokugam Hewage <chamin.lokugamhewage@ucdconnect.ie> |
| 11:30am – 12:00pm | S40205 - Vision Paper: Emergence of an Autonomous Vehicle Secondary Data Market for Breakthrough Applications | Kevin Dick <kevin.dick@carleton.ca> |
| 12:00pm – 12:30pm | S40202 - Trajectory-User Linking Is Easier Than You Think | Alameen Najjar <alameen.najjar@rakuten.com> |
| 12:30pm – 1:00pm | S40203 - Systematic Analysis of Public Transit Data Availability in Canada | Kevin Dick <kevin.dick@carleton.ca> |
| 1:00pm – 1:30pm | S40209 - Thematic Geo-Density Heatmapping for Walking Tourism Analytics using Semi-Ready GPS Trajectories | Iori Sasaki <sasaki@lab.akita-u.info> |
| 1:30pm – 1:45pm | **Closing Remarks** | |

## The 1st International Workshop on Neural Network Models for Brain Connectome Analysis (BrainNN2022)
*WorkshopChairs: Carl Yang, Xiaoxiao Li, Qingyu Zhao*

| Time | Title | Presenter/Author |
|---|---|---|
| 8:00-8:10 | Introduction and Opening Remarks | Dr. Carl Yang (no paper) |
| 8:10-8:35 | Invited Talk 1 | Dr. Yize Zhao (S41214) |
| 8:35-9:00 | Invited Talk 2 | Dr. Liang Zhan (S41211) |
| 9:00-10:00 | Spotlight Paper Presentations | Zijian Dong (S41201), Ethan Young (S41206), Xuan Kan (S41207), Alexis Li (S41208) |
| 10:00-10:20 | Coffee Break | |
| 10:20-10:45 | Invited Talk 3 | Dr. Archana Venkataraman (S41209) |
| 10:45-11:10 | Invited Talk 4 | Dr. Ying Guo (S41213) |
| 11:10-11:35 | Invited Talk 5 | Dr. Lifang He (S41212) |
| 11:35-12:00 | Invited Talk 6 | Dr. Ehsan Adeli (S41210) |
| 12:00-12:10 | Closing Remarks | |

## 1st International Workshop on Responsible AI and Data Ethics (RAIDE 2022) - Online
*Jonathan Boardman, Ying Xie*

| Time (JST) | Title | Presenter/Author |
|---|---|---|
| 8:30AM – 9:20AM | *Invited Talk - Why We Need to Formalize a Practically Oriented Sub-Discipline of AI and Data Ethics* | Jonathan Boardman |
| 9:30AM-9:50AM | S42201 *Humans as Mitigators of Biases in Risk Prediction via Field Studies* | Arul Mishra, Himanshu Mishra, and Bei Wang |
| 9:50AM-10:10AM | S42202 *Towards Implementing Responsible AI* | Conrad Sanderson, Qinghua Lu, David Douglas, Xiwei Xu, Liming Zhu, and Jon Whittle |
| 10:10AM-10:30AM | S42203 *Prioritizing Policies for Furthering Responsible Artificial Intelligence in the United States* | Emily Hadley |
| 10:50AM-11:10AM | S42204 *Siri, how long should this offender stay in prison?" Considerations about the use of algorithms for judgments in criminal proceedings* | Anna-Katharina Dhungel |
| 11:10AM-11:30AM | S42205 *Standardization on Bias in Artificial Intelligence as Industry Support* | Ewelina Szczekocka, Christèle Tarnec, and Janusz Pieczerak |
| 11:30AM-11:50AM | S42206 *The Analysis and Development of an XAI Process on Feature Contribution Explanation* | Jun Huang, Zerui Wang, Ding Li, and Yan Liu |
| | Break | |
| 12:20PM-12:40PM | S42207 *Entity Matching with AUC-Based Fairness* | Soudeh Nilforoushan, Qianfan Wu, and Mostafa Milani |
| 12:40PM-1:00PM | S42209 *Stereotype and Categorical Bias Evaluation via Differential Cosine Bias Measure* | Sudhashree Sayenju, Ramazan Aygun, Bill Franks, Sereres Johnston, George Lee, and Girish Modgil |
| 1:00PM-1:20PM | S42212 *Integrated Gradients is a Nonlinear Generalization of the Industry Standard Approach to Variable Attribution for Credit Risk Models* | Jonathan Boardman, Md Shafiul Alam, Xiao Huang, and Ying Xie |
| 1:20PM-1:40PM | S42210 *Applications of Integrated Gradients in Credit Risk Modeling* | Md Shafiul Alam, Jonathan Boardman, Xiao Huang, and Matthew Turner |
| 1:40PM-2:00PM | S42211 *An Application of Localized Model Explainability: Identifying Key Disparities in Social Determinants of Health in Food Deserts* | Md Shafiul Alam, Namazbai Ishmakhametov, Ying Xie, and Sumit Chakravarty |

| Time | Title | Presenter/Author |
|---|---|---|
| 2:00PM-2:20PM | BigD477 *Feature Integration Strategies for Multilingual Fake News Classification* | Jędrzej Kozal, Michał Leś, Paweł Zyblewski, Paweł Ksieniewicz, and Michał Woźniak |
| 2:20PM-2:40PM | BigD753 *Don't blindly use data. Towards a data statement for computational financial research.* | Stacey Taylor and Vlado Keselj |
| | **Closing Remarks** | |

## Creating Big Data for morals, ethics, and values in natural language in an automatically inferable format
### WorkshopChairs: Keisuke Nakamura & Toshiaki Kurokawa

| Time | Title | Presenter/Author |
|---|---|---|
| 9:00-9:25 | *Ethical Issues in Automatic Dialogue Generation for Non-Player Characters in Digital Games* #S44205 | Yusuke Mori, SQUARE ENIX AI & ARTS Alchemy Co., Ltd., and Youichiro Miyake, SQUARE ENIX |
| 9:25-9:50 | *Legal Artificial Assistance Agent to Assist Refugees* #S44207 | Alejandro T Gomez-Briones, Rappi Inc, Leslie G Gonzalez-Cardona, Jabil Inc, and Jovan D Rebolledo-Mendez, Exponential Japan |
| 9:50-10:00 | **Coffee Break** | |
| 10:00-10:25 | *Twitter Accounts Suggestion: Pipeline Technique SpaCy Entity Recognition* #S44211 | Shabbab Algamdi, Abduallah Albanyan, Sayed Shah, and Zeenat Tariq, University of North Texas |
| 10.25-10:50 | *General Framework for Fast Implementation for Conversational Artificial Assistance for Applications to Enable Global Impact as Tech-for-Good in Law* #S44208 | Alejandro T Gomez-Briones, Rappi Inc, Leslie G Gonzalez-Cardona, Jabil Inc, and Jovan D Rebolledo-Mendez, Exponential Japan |
| 10:50-11:00 | **Coffee Break** | |
| 11:00-11:25 | *Some Issues on Empirical Experiments and Ethical Rules for Knowledge Engineering Research in Interdisciplinary Fields* #S44212 | Munehiko Sasajima, University of Hyogo |
| 11.25-11:50 | *An Automatically Inferable Format for Creating Big Data of Morals, Ethics and ORDER OF VALUES Written Almost in Natural Language* #S44210 | Keisuke Nakamura, and Tatsuyoshi Ando, Knowrel System, Inc. |
| 11.50-12:15 | *A TABOO-NOT in OPEN World Assumption for A Natural Language based Logic Programming* #S44204 | Keisuke Nakamura, and Tatsuyoshi Ando, Knowrel System, Inc. |
| **12:15-12:20** | **Closing Remarks** | |

## 11th Workshop on Scalable Cloud Data Management
### WorkshopChairs: Felix Gessert, Wolfram Wingerath, Benjamin Wollmer, Norbert Ritter

| Time | Title | Presenter/Author |
|---|---|---|
| **3:00 - 3:10 pm** | **Opening Remarks** | |
| 3:10 - 4:10 pm | **Keynote** Benchmarking for Big Data Integration: Open Challenges & Research Directions | Fabian Panse (University of Hamburg) |

| | | |
|---|---|---|
| 4:10 - 5:10 pm | Metadata Caching in Presto: Towards Fast Data Processing (S04204) | Beinan Wang, Chunxu Tang, Rongrong Zhong, Bin Fan, Yi Wang, Jasmine Wang, Shouwei Chen, Bowen Ding and Lu Zhang (Alluxio, Inc; Twitter, Inc; Celonis, Inc) |
| | Towards Advanced Monitoring for Scientific Workflows (S04202) | Jonathan Bader, Joel Witzke, Soeren Becker, Ansgar Lößer, Fabian Lehmann, Leon Doehler, Anh Duc Vu and Odej Kao (Technische Universität Berlin; Zuse Institute Berlin; TU Darmstadt; Humboldt-Universitat zu Berlin) |
| **5:10 - 5:30 pm** | **Coffee Break** | |
| 5:30 - 7:00 pm | Hybrid Cloud and HPC Approach to High-Performance Dataframes (S04206) | Kaiying Shan, Niranda Perera, Damitha Lenadora, Tianle Zhong, Arup Sarker, Supun Kamburugamuve, Thejaka Amila Kanewela, Chathura Widanage, and Geoffrey Fox (Luddy School of Informatics, Computing, and Engineering; University of Illinois Urbana-Champaign; Indiana University Alumni) |
| | FineCloud: Fine-grained Cloud Service Advisory using Machine Learning (S04205) | Yasemin Orhun, Yigit Istanbullu and Ismail Ari (Ozyegin University) |
| | A Deep Learning-based Prefetching Approach to Enable Scalability for Data-intensive Applications (S04203) | Tolga Büyüktanır and Mehmet S. Aktas (Yildiz Technical University) |
| **7:00 - 7:15 pm** | **Closing Remarks** | |

# Special Sessions

## 5th Special Session on HealthCare Data <mark>(ONLINE)</mark>

| Time | Title | Presenter/Author |
|---|---|---|
| **Day 1: Saturday, December 17, 2022** | | |
| **Time** | **Title** | **Presenter/Author** |
| **14:00-14:10** | *Session Opening* | |
| **14:10-14:30** | SP01201 - Evolving Perceptions of Mental Health on Social Media and their Medical Impacts | Kenna Reagan, Aparna Varde, and Lei Xie |
| **14:30-14:50** | SP01204 - DPPIN: A Biological Repository of Dynamic Protein-Protein Interaction Network Data | Dongqi Fu and Jingrui He |
| **14:50-15:10** | SP01205 - Is Unimodal Bias Always Bad for Visual Question Answering? A Medical Domain Study with Dynamic Attention | Zhongtian Sun, Anoushka Harit, Alexandra I. Cristea, Jialin Yu, Noura Al Moubayed, and Lei Shi |
| **15:10-15:30** | SP01206 - Deep Pneumonia: Attention-Based Contrastive Learning for Class-Imbalanced Pneumonia Lesion Recognition in Chest X-rays | Xinxu Wei, Xiangke Niu, Xianshi Zhang, and Yongjie Li |
| **15:30-15:50** | SP01209 - Explainability Using SHAP for Epileptic Seizure Recognition | Simone Ludwig |
| **15:50-16:10** | SP01211 - Semantic Dietary Personalization for Chronic Disease Management | Changhyub Woo, Byunghyun Kim, Jaehun Lee, Buhwan Jeong, Jinwoo Song, Carlo Allocca, and Rohit Ail |
| **16:10-16:20** | *Break* | |
| **16:20-16:40** | SP01212 - Prototyping the Problem Oriented Medical Record for Connected Health Based on TypeGraphQL | Sabah Mohammed, Jinan Fiaidhi, and Darien Sawyer |
| **16:40-17:00** | SP01213 - Thick Data Techniques for Identifying Abnormality in Video Frames for Wireless Capsule Endoscopy | Jinan Fiaidhi, Sabah Mohammed, and Petros Zezos |
| **17:00-17:20** | SP012014 - Predicting Taxonomic Identity and Genetic Composition of Codon Usage Bias Levels Using Deep Learning Models | Lennart Buhl and Sayantica Pattanayak |
| **17:20-17:40** | SP01215 - DeePaGait: Motor Assessment of Parkinson's Disease Using a Multi-Layer 1D Convolutional Neural Network on Smartphone Gait Data | Hamza Abujrida, Emmanuel Agu, and Kaveh Pahlavan |
| **17:40-18:00** | SP01216 - Explainable Prediction of the Severity of COVID-19 Outbreak for US Counties | Shailik Sarkar, Abdulaziz Alhamadani, and Chang-Tien Lu |
| **18:00-18:10** | *Session Closing* | |
| **Day 2: Sunday, December 18, 2022** | | |
| **Time** | **Title** | **Presenter/Author** |

| 14:00-14:10 | Session Opening | |
|---|---|---|
| 14:10-14:30 | SP012017 - Building a Patient-Centered Virtual Hospital Ecosystem Using Both Access Control and CNN-Based Models | Sara A. Alsalamah, Shada AlSalamah, Walaa N. Ismail, Hessah Alsalamah, and Chang-Tien Lu |
| 14:30-14:50 | SP01218 - Anonymizing Periodical Releases of SRS Data by Fusinging Differential Privacy | Yi-Yuang Wu, Zhi-Xun Shen, and Wen-Yang Lin |
| 14:50-15:10 | SP01219 - HAR-CTGAN: A Mobile Sensor Data Generation Tool for Human Activity Recognition | Joshua DeOliveira, Walter Gerych, Aruzhan Koshkarova, Elke Rundensteiner, and Emmanuel Agu |
| 15:10-15:30 | SP01220 - Deep Video Action Recognition Models for Assessing Cardiac Function from Echocardiograms | Abdulsalam Almadani, Abhishek Shivdeo, Emmanuel Agu, and Jacques Kpodonu |
| 15:30-15:50 | SP01221 - Oversampling techniques for predicting COVID-19 patient length of stay | Zachariah Farahany, Jiawei Wu, K M Sajjadul Islam, and Praveen Madiraju |
| 15:50-16:10 | SP01222 - LSTM-based Pulmonary Air Leak Forecasting for Chest Tube Management | Roberto Corizzo, Rodrigo Yepez-Lopez, Sebastien Gilbert, and Nathalie Japkowicz |
| 16:10-16:20 | Break | |
| 16:20-16:40 | SP01223 - CCS Explorer: Relevance Prediction, Extractive Summarization, and Named Entity Recognition from Clinical Cohort Studies | Irfan Al-Hussaini, Davi Nakajima An, Albert J. Lee, Sarah Bi, and Cassie S. Mitchell |
| 16:40-17:00 | SP01224 - Typhon: Parallel Transfer on Heterogeneous Datasets for Cancer Detection in Computer-Aided Diagnosis | Giuseppe Cuccu, Christophe Broillet, Carolin Reischauer, Harriet Thoeny, and Philippe Cudre-Mauroux |
| 17:00-17:20 | SP01225 - Data-Driven Estimation of Effectiveness of COVID-19 Non-pharmaceutical Intervention Policies | yash mahajan, Sheikh Rabiul Islam, Mohammad Ruhul Amin, and Shubhra (``Santu&#039;&#039;) Karmaker |
| 17:20-17:40 | SP01226 - DOD-Explainer: Explainable Drug Overdose Deaths Predictor from Crime and Socioeconomic Data | Abdulaziz Alhamadani, Shailik Sarkar, Lulwah Alkulaib, and Chang-Tien Lu |
| 17:40-18:00 | BigD314 - Swin-TCNet: A vision transformer-based method for automated chromosomal object detection | Chih-En Kuo, Chien-Hsing Lu, and Li-Yuan Huang |
| 18:00-18:20 | SP01208 - Deep Reinforcement Learning for Bandit Arm Localization | Wenbin Du, Huaqing Jin, Chao Yu, and Guosheng Yin |
| 18:20-18:30 | Session Closing | |
| **Day 3: Monday, December 19, 2022** | | |
| Time | Title | Presenter/Author |
| 14:00-14:10 | Session Opening | |

| Time | Title | Presenter/Author |
|---|---|---|
| 14:10-14:30 | BigD549 - Preserving Cluster Features in Imputing High Dimensional Data with Extensive Missing Rate | Chih Lai, Carolin Poschen, Lisa Steinheuer, and Jorg Hackermuller |
| 14:30-14:50 | BigD593 - An Automatic Sleep Arousal Detection Method by Enhancing U-Net with Spatial-channel Attention | Yen-Ning Su and Chih-En Kuo |
| 14:50-15:10 | BigD691 - Pseudo slicer on three dimensional brain tumor segmentation | Mohammad Iqbal, Fitria Urbach, Hsing-Kuo Pao, Anggraini Dwi Sensusiati, Nurul Hidayat, and Imam Mukhlash |
| 15:10-15:30 | BigD705 - Evaluating COVID-19 Sequence Data Using Nearest-Neighbors Based Network Model | Sarwan Ali |
| 15:30-15:50 | BigD779 - VF2VF: Improving Precision while Maintaining Accuracy | Zhenyu Zhang and Haogang Zhu |
| 15:50-16:00 | Session Closing | |

# 2022 Special Session on Machine Learning on Big Data

| Day 1: Monday, December 19, 2022 - ONSITE | | |
|---|---|---|
| Time | Title | Presenter/Author |
| 10:00 – 10:30 | Coffee Break | |
| 10:30-10:50 | SP02203 Multi-class Skin Cancer Classification Architecture Based on Deep Convolutional Neural Network | Mst Shapna Akter, Hossain Shahriar, Sweta Sneha, Alfredo Cuzzocrea, and Carson K. Leung |
| 10:50-11:10 | SP02205 DVGAN: Stabilize Wasserstein GAN training for time-domain Gravitational Wave physics | Tom Dooney, Stefano Bromuri, and Lyana Curier |
| 11:10-11:30 | SP02208 Causal Discovery for Feature Selection in Physical Process-Based Hydrological Systems | Paras Sheth, Ting Liu, Durmus Doner, Yuhang Wei, Qi Deng, Rebecca Muenich, John Sabo, K. Selcuk Candan, and Huan Liu |
| 11:30-11:50 | SP02209 STCD: A Spatio-Temporal Causal Discovery Framework for Hydrological Systems | Paras Sheth, Reepal Shah, John Sabo, K. Selcuk Candan, and Huan Liu |
| 11:50-12:10 | SP02211 GAPS: Generality and Precision with Shapley Attribution | Brian Daley, Qudrat Ratul, Edoardo Serra, and Alfredo Cuzzocrea |
| 12:10-12:30 | | |
| 12:10-12:30 | SP02218 Handwritten Word Recognition using Deep Learning Approach: A Novel Way of Generating Handwritten Words | Mst Shapna Akter, Hossain Shahriar, and Alfredo Cuzzocrea |
| 12:30 – 14:00 | Lunch Break | |
| 14:00-14:20 | SP02221 Learning Latent Road Correlations from Trajectories | Zheng Dong, Quanjun Chen, Renhe Jiang, |

| Time | Title | Presenter/Author |
|---|---|---|
| | | Huanchen Wang, Xuan Song, and Hao Tian |
| 14:20-14:40 | **SP02222** A Crowd Source System for Big Data Analytics for YouTube | Jagan Mohan Reddy Danda, Abhishek Attuluri, Hossain Shahriar, Alfredo Cuzzocrea, and Nazmus Sakib |
| 14:40-15:00 | **SP02223** An Empirical Study of Domain Adaptation: Are We Really Learning Transferable Representations? | Nicholas Josselyn, Biao Yin, Ziming Zhang, and Elke Rundensteiner |
| 15:00-15:20 | **N237** Efficient Comprehensive Element Identification in Large Scale Spectral Analysis with Interpretable | Xiang Xie and Wilhelm Stork |
| 15:20-15:40 | **BigD306** OSP-Class: Open Set Pseudo-labeling with Noise Robust Training for Text Classification | Dohyung Kim, Jahwan Koo, and Ung-Mo Kim |
| 15:40-16:00 | **BigD555** Badminton stroke classification based on accelerometer data: from individual to generalized models | Daniel Peralta, Ben Van Herbruggen, Jaron Fontaine, Wout Debyser, Jorg Wieme, and Eli De Poorter |
| 16:00 – 16:30 | **Coffee Break** | |

## Day 1: Monday, December 19, 2022 - ONLINE

| Time | Title | Presenter/Author |
|---|---|---|
| 16:30-16:50 | **SP02206** Graph Clustering using Node Embeddings: An Empirical Study [Short Paper] | Mahdi Ghanbari and Mostafa Haghir Chehreghani |
| 16:50-17:10 | **SP02207** Autonomous Drug Design with Multi-armed Bandits | Hampus Gummesson Svensson, Esben Bjerrum, Christian Tyrchan, Ola Engkvist, and Morteza Haghir |
| 17:10-17:30 | **SP02210** Distributional Shift Adaptation using Domain-Specific Features | Anique Tahir, Lu Cheng, Ruocheng Guo, and Huan Liu |
| 17:30-17:50 | **SP02216** A Novel Machine Learning Based Framework for Bridge Condition Analysis | Mohammad Masum, Nafisa Anjum, Hossain Shahriar, Nazmus Sakib, Maria Valero, M.A. Karim, Akond Rahman, Fan Wu, and Alfredo Cuzzocrea |
| 17:50-18:10 | **SP02220** Multi-criteria Rating and Review based Recommendation Model | Emrul Hasan, Chen Ding, and Alfredo Cuzzocrea |
| 18:10-18:30 | **SP02224** Distributed Heterogeneous Transfer Learning for Link Prediction in the Positive Unlabeled Setting | Paolo Mignone, Gianvito Pio, and Michelangelo Ceci |
| 18:30-18:50 | **SP02225** Deep Federated Anomaly Detection for Multivariate Time Series Data | Wei Zhu, Dongjin Song, Yuncong Chen, Wei Cheng, Bo Zong, Takehiko Mizoguchi, Cristian Lumezanu, Haifeng Chen, and Jiebo Luo |

## Day 2: Tuesday, December 20, 2022 - ONSITE

| Time | Title | Presenter/Author |
|---|---|---|
| 10:00 – 10:30 | **Coffee Break** | |
| 10:30-10:50 | **BigD560** Density-Based Prototypical Contrastive Learning on Visual Representations | Yueh-Lin Chung and Hsing-Kuo Pao |

| Time | Title | Presenter/Author |
|---|---|---|
| 10:50-11:10 | **BigD664** Variational Factorization Machines for Preference Elicitation in Large-Scale Recommender Systems | Jill-Jenn Vie, Tomas Rigaux, and Hisashi Kashima |
| 11:10-11:30 | **BigD698** HITS-GNN: A Simplified Propagation Scheme for Graph Neural Networks | Mehak Khan, Gustavo B. M. Mello, Paal Engelstad, Laurence Habib, and Anis Yazidi |
| 11:30-11:50 | **BigD729** Augmentation in the Embedding space for Contrastive Representation Distillation | Yi-Chieh Cheng and Hsing-Kuo Pao |
| 11:50-12:10 | **BigD836** Federated Learning for Autoencoder-based Condition Monitoring in the Industrial Internet of Things | Soeren Becker, Kevin Styp-Rekowski, Oliver Vincent Leon Stoll, and Odej Kao |
| 12:10-12:30 | **BigD844** AsyncFLEO: Asynchronous Federated Learning for LEO Satellite Constellations with High-Altitude-Platforms | Mohamed Elmahallawy, and Tony T. Luo |
| 12:30 – 14:00 | **Lunch Break** | |
| **Day 2: Tuesday, December 20, 2022 - ONLINE** | | |
| Time | Title | Presenter/Author |
| 16:30-16:50 | **SP02226** Leveraging Knowledge Graph Embeddings to Disambiguate Author Names in Scientific Data | Laura Rettig, Kurt Baumann, Sebastian Sigloch, and Philippe Cudré-Mauroux |
| 16:50-17:10 | **SP02230** On Designing Day Ahead and Same Day Ridership Level Prediction Models for City-Scale Transit Networks Using Noisy APC Data | Jose Paolo Talusan, Ayan Mukhopadhyay, Dan Freudberg, and Abhishek Dubey |
| 17:10-17:30 | **BigD446** More is Less - Byte-quantized models are faster than bit-quantized models on the edge | Pengfei Zhang, Chenxia Han, and Eric Lo |
| 17:30-17:50 | **BigD495** Stable Topological Feature Vectors via Hermite Function Expansion on Persistence Curves | Yu-Min Chung |
| 17:50-18:10 | **BigD545** Scalable Learning with Incremental Probabilistic PCA | Boshi Wang and Adrian Barbu |
| 18:10-18:30 | **BigD748** Analysing Longitudinal Social Science Questionnaires: Topic modelling with BERT-based Embeddings | Vida Sharifian-Attar, Suparna De, Sanaz Jabbari, Jenny Li, Harry Moss, and Jon Johnson |
| 18:30-18:50 | **SP02214** Landslide Likelihood Prediction using Machine Learning Algorithms | Vasundhara Acharya, Anindita Ghosh, Inwon Kang, Thilanka Munasinghe, and Binita KC |

# 2022 Special Session on Privacy and Security of Big Data

| Day 1: Monday, December 19, 2022 - ONSITE | | |
|---|---|---|
| Time | Title | Presenter/Author |
| 10:00 – 10:30 | **Coffee Break** | |
| 10:30-10:50 | **SP03201** Homomorphic Encryption-Friendly Privacy-Preserving Partitioning Algorithm for Differential Privacy | Shojiro Ushiyama, Tsubasa Takahashi, Masashi Kudo, and Hayato Yamana |

| Time | Title | Presenter/Author |
|---|---|---|
| 10:50-11:10 | **SP03202** Anomaly Detection in Cybersecurity Events a Through Graph Neural Network and Transformer Based Model: A Case Study with BETH Dataset | Bishal Lakha, Sara Mount, Edoardo Serra, and Alfredo Cuzzocrea |
| 11:10-11:30 | **SP03203** Adaptive Attacks and Targeted Fingerprinting of Relational Data | Tanja Sarcevic, Rudolf Mayer, and Andreas Rauber |
| 11:30-11:50 | **SP03204** Efficient Bayesian Network Construction for Increased Privacy on Synthetic Data | Markus Hittmeir, Rudolf Mayer, and Andreas Ekelhart |
| 11:50-12:10 | **SP03209** An efficient approach for anonymising heterogeneous graph structures | Guillermo Alaman Requena, Andreas Ekelhart, and Rudolf Mayer |
| 12:10-12:30 | **SP03210** Deep Learning Approach for Classifying the Aggressive Comments on Social Media: Machine Translated Data | Mst Shapna Akter, Hossain Shahriar, and Alfredo Cuzzocrea |
| 12:10-12:30 | **BigD725** An Unified Noise and Watermark Removal from Information Bottleneck-based Modeling | Han-Juan Huang, Sheng-Wen Chen, and Hsing-Kuo Pao |
| 12:30 – 14:00 | **Lunch Break** | |
| 14:00-14:20 | **BigD727** Differential Privacy-inspired Federated Split Learning for Anomaly Detection | Chi-Chen Liao and Hsing-Kuo Pao |
| 14:20-14:40 | **BigD764** Mixers Detection in bitcoin network: a step towards detecting money laundering in crypto-currencies | Muhammad Mazhar Ullah Rathore, Sushil Chaurasia, and Dhirendra Shukla |
| 14:40-15:00 | **BigD827** Permutation-Invariant Tabular Data Synthesis | Yujin Zhu, Zilong Zhao, Robert Birke, and Lydia Y. Chen |
| 15:00-15:20 | **BigD839** A Crypto-Assisted Approach for Publishing Graph Statistics with Node Local Differential Privacy | Shang Liu, Yang Cao, Takao Murakami, and Masatoshi Yoshikawa |
| 15:20-15:40 | | |
| 15:40-16:00 | | |
| 16:00 – 16:30 | **Coffee Break** | |
| **Day 1: Monday, December 19, 2022 - ONLINE** | |
| **Time** | **Title** | **Presenter/Author** |
| 16:30-16:50 | **SP03206** Software Supply Chain Vulnerabilities Detection in Source Code: Performance Comparison between Traditional and Quantum Machine Learning Algorithms | Mst Shapna Akter, MD Jobair Hossain Faruk, Nafisa Anjum, Mohammad Masum, Hossain Shahriar, Nazmus Sakib, Maria Valero, Akond Rahman, Fan Wu, and Alfredo Cuzzocrea |
| 16:50-17:10 | **SP03207** Authentic Learning of Machine Learning in Cybersecurity with Portable Hands-on Labware: Neural Network Algorithms for Network Denial of Service (DOS) Detection | MD Jobair Hossain Faruk, Hossain Shahriar, Kai Qian, Dan Lo, Michael Whitman, Alfredo Cuzzocrea, and Fan Wu |
| 17:10-17:30 | **SP03208** Predicting and Analyzing Privacy Settings and Categories for Posts on Social Media | Hsin-Yu Chen and Cheng-Te Li |
| 17:30-17:50 | **SP03211** Explaining IoT Attacks: An Effective and Efficient Semi-Supervised Learning Framework | Giuseppe Cascavilla, Reinier Zwart, Damian Tamburri, and Alfredo Cuzzocrea |

| Time | Title | Presenter/Author |
|---|---|---|
| 17:50-18:10 | **SP03212** Privacy and security of mobile users in smart cities: an overview | Sabrina Ichou, Slimane Hammoudi, Amel Benna, Abdelkrim Meziane, and Alfredo Cuzzocrea |
| 18:10-18:30 | **SP03213** On using Node Indices and Their Correlations for Fake Account Detection | Sara Asghari and Mostafa Haghir Chehreghani |
| 18:30-18:50 | **N294** HateNet: A Graph Convolutional Network Approach to Hate Speech Detection | Charles Duong, Lei Zhang, and Chang-Tien Lu |
| **Day 2: Tuesday, December 20, 2022 - ONLINE** | | |
| **Time** | **Title** | **Presenter/Author** |
| 10:00 – 10:30 | **Coffee Break** | |
| 10:30-10:50 | | |
| 10:50-11:10 | **BigD265** Differentially Private Federated Continual Learning with Heterogeneous Cohort Privacy | Ajesh Koyatan Chathoth, Clark P Necciai Jr, Abhyuday Jagannatha, and Stephen Lee |
| 11:10-11:30 | **BigD322** Exploring the Generalisability of Fake News Detection Models | Nathaniel Hoy and Theodora Koulouri |
| 11:30-11:50 | **BigD361** One Stone, Three Birds: Finer-Grained Encryption with Apache Parquet @ Large Scale | Xinli. Shang, Pavi Subenderan, Mohammad Islam, and Jianchun Xu |
| 11:50-12:10 | **BigD386** Twitter Bot Detection Through Unsupervised Machine Learning | Jeremy Wu, Eric Teng, and Ziyue Cao |
| 12:10-12:30 | **BigD510** Diseases Prediction from Officially Anonymized Medical and Healthcare Big Data | Hiroaki Kikuchi, Satoshi Ito, Kazuki Ikegami, and Shota Shindo |
| 12:30 – 14:00 | **Lunch Break** | |
| 16:30-16:50 | **BigD544** Efficient Inference on Convolutional Neural Networks by Image Difficulty Prediction | Yu-Jen Chang, Ding-Yong Hong, Pangfeng Liu, and Jan-Jan Wu |
| 16:50-17:10 | **BigD623** Federated XGBoost on Sample-Wise Non-IID Data | Katelinh Jones, Yuya Ong, Yi Zhou, and Nathalie Baracaldo |
| 17:10-17:30 | **BigD631** Low-rank Defenses Against Adversarial Attacks in Recommender Systems | Negin Entezari and Evangelos Papalexakis |
| 17:30-17:50 | **BigD708** A Novel Video Copy Subsequence Detection and Location Method | Xie Yuxiang, Yan Jie, Luan Xidao, Zhang Lili, Zhao Yancheng, and Wei Yingmei |
| 17:50-18:10 | **BigD722** Federated Local Differential Privacy for Recommender Systems | Manoj Reddy Dareddy and Junghoo Cho |
| 18:10-18:30 | **BigD744** Analyzing and Defending against Membership Inference Attacks in Natural Language Processing | Yijue Wang, Nuo Xu, Shaoyi Huang, Kaleel Mahmood, Dan Guo, Caiwen Ding, Wujie Wen, and Sanguthevar Rajasekaran |
| 18:30-18:50 | **BigD757** Robust Stochastic Bandit algorithms to defend against Oracle attack using Sample Dropout | Jayanth Yetukuri and Yang Liu |

# 8th Special Session on Intelligent Data Mining <mark>(ONLINE)</mark>

| Time | Paper ID | Presenter/Author |
|---|---|---|
| **Day 1: Saturday, December 17, 2022** | | |
| 07:00am-08:00am | | **Registration** |
| 08:00am-08:10am | | **Session Keynote Speech**<br>Uraz Yavanoglu, PhD |
| 08:10am-08:20am | SP04201 | Joshua Lampkins, Darren Chan, Alan Perry, Sasha Strelnikoff, Jiejun Xu, and Alireza Esna Ashari<br>Multimodal Road Sign Interpretation for Autonomous Vehicles |
| 08:20am-08:30am | SP04202 | Salim Sazzed<br>The Dynamics of Ukraine-Russian Conflict through the Lens of Demographically Diverse Twitter Data |
| 08:30am-08:40am | SP04204 | Girmaw Abebe Tadesse, William Ogallo, Celia Cintas, and Skyler Speakman<br>Model-free feature selection to facilitate automatic discovery of divergent subgroups in tabular data |
| 08:40am-08:50am | SP04205 | Rafael Hidalgo, Anthony DeVito, Nesreen Salah, Aparna Varde, and Robert Meredith<br>Inferring Phylogenetic Relationships using the Smith-Waterman Algorithm and Hierarchical Clustering |
| 08:50am-09:00am | SP04208 | Yuto Kosugi, Itsuki Matsunaga, Hangli Ge, Takashi Michikata, and Noboru Koshizuka<br>Traffic Congestion Prediction Using Toll and Route Search Log Data |
| 09:00am-09:10am | SP04209 | Djellel Difallah, Diego Saez-Trumper, Eriq Augustine, Robert West, and Leila Zia<br>Crosslingual Section Title Alignment in Wikipedia |
| 09:10am-09:20am | SP04210 | Natalia Ostapuk, Djellel Difallah, and Philippe Cudre-Mauroux<br>ParaGraph: Mapping Wikidata Tail Entities to Wikipedia Paragraphs |
| 09:20am-09:30am | SP04211 | Lars Hillebrand, Tobias Deußer, Tim Dilmaghani, Bernd Kliem, Rüdiger Loitz, Christian Bauckhage, and Rafet Sifa<br>Towards automating Numerical Consistency Checks in Financial Reports |
| 09:30am-09:40am | SP04214 | Yufei Zhao, Chao Huang, and Jiebo Luo<br>How to Prepare for the Next Pandemic – Investigation of Correlation Between Food Prices and COVID-19 From Global and Local Perspectives |
| 09:40am-10:00am | | **Break** |
| 10:00am-10:10am | SP04215 | Leah Ding, Roberto Corizzo, Colin Bellinger, Nancy Ching, Spencer Login, Rodrigo Yepez-Lopez, Jie Gong, and Dong L. Wu<br>Imbalanced Multi-layer Cloud Classification with Advanced Baseline Imager (ABI) and CloudSat/CALIPSO Data |
| 10:10am-10:20am | SP04216 | Ragini Kihlman and Maria Fasli<br>Improving the co-training algorithm to enhance semi-supervised learning results |

| Time | Paper ID | Title |
|---|---|---|
| 10:20am-10:30am | SP04217 | Sajid Hussain Rafi Ahamed, Shahid Shakil, Hanjia Lyu, Xinping Zhang, and Jiebo Luo<br>Doctors vs. Nurses: Understanding the Great Divide in Vaccine Hesitancy among Healthcare Workers |
| 10:30am-10:40am | SP04218 | Jack Tacchi, Chiara Boldrini, Andrea Passarella, and Marco Conti<br>Signed Ego Network Model and its Application to Twitter |
| 10:40am-10:50am | SP04212 | Perin Ünal, Özlem Albayrak, Meerim Kubatova, Bilgin U. Deveci, Ege Çırakman, İpek Koçal, and A. Murat Özbayoğlu<br>A Big Data Application in Manufacturing Industry- Computer Vision to Detect Defects on Bearings |
| 10:50am-11:00am | SP04220 | Seyithan Temel, Emre Ummak, Abdülkadir Tokgöz, Furkan Işık, Özlem Albayrak, Perin Ünal, and Murat Özbayoğlu<br>Control System Design and Implementation Based on Big Data and Ontology |
| 11:00am-11:10am | BigD203 | Ruijie Tian, Weishi Zhang, Fei Wang, and Junting Xiong<br>A Context-Aware Method for Indexing Large-Scale SpatioTemporal Data |
| 11:10am-11:20am | BigD226 | Sukhwan Jung and Aviv Segev<br>Semantic Similarity Analysis between Future Topics and Their Neighbors in Topic Networks for Network-based Topic Evolution |
| 11:20am-11:30am | BigD242 | Ao Wang, Yongchen Yao, Yuhan Yao, and Jinpeng Chen<br>Nomaber: A Novel Framework For Multi-Type Abnormal Behavior Recognition |
| 11:30am-11:40am | BigD258 | Cathy Zhang and M Omair Shafiq<br>A Real-time, Scalable Monitoring and User Analytics Solution for Microservices-based Software Applications |
| 11:40am-11:50am | BigD325 | Yifan Xu, Minghao Zhu, Junli Wang, and Chungang Yan<br>LAMPT: LAbel Mask-Predict Transformer for Extreme Multi-label Text Classification |
| 11:50am-12:00pm | BigD383 | Donald Bleyl and Elham Buxton<br>Emotion Recognition on StackOverflow Posts Using Bert |
| 12:00pm-12:10pm | BigD469 | Zhou Huang and Feng Zhou<br>An Approximation Method for Querying Similar Large Graphs |
| 12:10pm-1:00pm | | **Lunch Break** |
| 1:00pm-1:10pm | BigD515 | Tomomi Yamazaki, Takahiro Miura, and Ichiro Sakata<br>Big data analysis reveals an emerging change in academia-industry collaborations in the era of digital convergence |
| 1:10pm-1:20pm | BigD622 | Jabir Alshehabi Al-Ani and Maria Fasli,<br>*Probabilistic Named Entity Recognition for non-standard format entities using co-occurrence word embeddings* |
| 1:20pm-1:30pm | BigD629 | Sarwan Ali, Muhammad Ahmad, Umair ul Hassan, Muhammad Asad Khan, Shafiq Alam, and Imdadullah Khan<br>Efficient Data Analytics on Augmented Similarity Triplets |

| Time | | |
|---|---|---|
| 1:30pm-1:40pm | BigD639 | Lizhuang Sun, Wenhui Hu, Kehui Xu, Yuhang Chen, Qidong Sun, and Jiaxi Wang<br>ASRC:A Knowledge Graph Relation Construction Model based on Active Learning and Semantic Recognition |
| 1:40pm-1:50pm | BigD648 | Ngai Lam Ho and Kwan Hui Lim<br>POIBERT: A Transformer-based Model for the Tour Recommendation Problem |
| 1:50pm-2:00pm | BigD750 | Haris Mansoor, Sarwan Ali, Shafiq Alam, Muhammad Asad Khan, Umair ul Hassan, and Imdadullah Khan<br>Impact Of Missing Data Imputation On The Fairness And Accuracy Of Graph Node Classifiers |
| 02:00pm-02:10pm | BigD851 | Jenq-Haur Wang and Tzong-Ting Hsieh<br>Dot-Product Similarity based Centroid Clustering for Popular Topic Detection |
| 02:10pm-02:20pm | SP04219 | Mustafa Çataltaş, Semih Yumusak, and Kasim Oztoprak<br>Academic Graph: A Literature Review System |
| 02:20pm-02:30pm | SP04206 | Semih Yumusak, Mustafa Coban, Yasin Yilmaz, and Hüseyin Oktay Altun<br>Detecting Dangerous Maritime Refugee Migration Paths through Cell Phone Activities |
| 02:30pm-02:40pm | BigD610 | Manh-Phu Nguyen, Minh-Son Dao, and Thanh-Binh Nguyen<br>3D-STGPCN: 3D Spatio-Temporal Graph Point-wise Convolutional Network for Traffic Forecasting |
| 02:40pm-02:50pm | BigD671 | Khoa Anh Tran, Do-Van Nguyen, Minh-Son Dao, and Koji Zettsu<br>SplitDyn: Federated Split Neural Network for Distributed Edge AI Applications |
| 03:00pm | Session Closing | Session Closing |

# 8th Special Session on Information Granulation in Data Science and Scalable Computing (PART 1 ONLINE, PART 2 ONSITE)

| Day 1: Saturday, December 17, 2022 - ONLINE | | |
|---|---|---|
| **Time** | **Title** | **Presenter/Author** |
| **16:00-16:20** | Accelerating concept-dependent granulation technique using data decomposition | Radoslaw Cybulski and Piotr Artiemjew |
| **16:20-16:40** | Tensor-based Approach to Big Data Processing and Machine Learning | Maciej Bartoszuk, Jaroslaw Litwin, Mateusz Wnuk, and Dominik Slezak |
| **16:40-17:00** | Targeted Mining of Rare High-utility Patterns | Peifeng Zhang, Jiahui Chen, Shicheng Wan, and Wensheng Gan |
| **17:00-17:20** | Flexibly Mining Better Patterns | Gengsen Huang, Wensheng Gan, Long Li, and Tianlong Gu |
| **17:20-17:40** | Pattern Discovery with Utility Occupancy | Jiayi Sun, Wensheng Gan, Jerry Chun-Wei Lin, and Han-Chieh Chao |
| **17:40-18:00** | Dataset Generation Framework for Evaluation of IoT Linux Host-Based Intrusion Detection Systems | Blazej Adamczyk, Maksym Brzeczek, Marcin Michalak, Iwona Kostorz, Lukasz |

| Time | | Wawrowski, Marek Hermansa, Michal Czerwinski, and Antoni Jamiolkowski |
|---|---|---|
| **18:00-18:20** | Incremental Fuzzy Utility Mining with Tree Structure | Tzung-Pei Hong, Wei-Teng Hung, Wei-Ming Huang, and Yu-Chuan Tsai |
| **18:20-18:40** | Unified Temporal Erasable Itemset Mining with a Lower-Bound Strategy | Tzung-Pei Hong, Jia-Xiang Li, Yu-Chuan Tsai, and Wei-Ming Huang |
| **18:40-19:00** | Large Scale Windowed Matching | Mateusz Przyborowski, Krzysztof Ciebiera, and Krzysztof Stencel |
| <td colspan="3" align="center">**Day 2: Sunday, December 18, 2022 - ONSITE**</td> | | |
| **16:20-16:40** | EVEAL - Expected Variance Estimation for Active Learning | Daniel Kaluza, Andrzej Janusz, and Dominik Slezak |
| **16:40-17:00** | Adversarial Robustness of Uncertainty-Aware Deep Learning Decision-Making Systems | Lukasz Grad |
| **17:00-17:20** | Similarity Analysis of Order Trajectory for Hospital Management | Tomohiro Kimura, Shusaku Tsumoto, and Shoji Hirano |
| **17:20-17:40** | Patients' Movement Analysis using Hospital Information System | Tomohiro Kimura, Shusaku Tsumoto, and Shoji Hirano |
| **17:40-18:00** | Approximation of the expectation-maximization algorithm for Gaussian mixture models on big data | Mateusz Przyborowski and Dominik Slezak |
| **18:00-18:20** | Data management in training AI chatbot with personality based on granular computing | Piotr Podolski, Tomasz Ludziejewski, and Hung Son Nguyen |

# Special Session on Platform for DFFT (Data Free Flow with Trust) <mark>(ONSITE)</mark>

| Time | Title | Presenter/Author |
|---|---|---|
| <td colspan="3" align="center">*WorkshopChairs: Noboru Koshizuka, noboru@koshizuka-lab.org*</td> | | |
| **14:00~14:05** | <td colspan="2" align="center">**Opening Remarks**</td> | |
| 14:05~14:25 | SP06201: "Toward High-Resolution Regional Atmospheric Reanalysis for Japan: An Overview of the ClimCORE Project" | Hisashi Nakamura, Kenichi Kuma, Kazutoshi Onogi, Takafumi Miyasaka, Yasutaka Makihara, Junichi Ishida, and Makoto Iida |
| 14:25~14:45 | SP06202: "DATA-EX: Infrastructure for Cross-Domain Data Exchange Based on Federated Architecture" | Noboru Koshizuka and Hiroshi Mano |
| 14:45~15:05 | SP06203: "Data Leaves as Scenario-oriented Metadata for Data-driven Innovation on Trust" | Yukio Ohsawa, Kaira Sekiguchi, Tomohide Maekawa, Hiroki Yamaguchi, Yeon Hyuk Son, and Sae Kondo |
| 15:05~15:25 | SP06204: "Design for Data Structures : Data Unification and Federation with Wikibase" | Hiroki Uematsu, Phuc Nugyen, and Hideaki Takeda |
| 15:25~15:45 | BigD661: "The concept of data trading market with data usage rights and its approach to implement" | Hiroshi Mano and Keita Saito |
| 15:45~15:55 | <td colspan="2" align="center">Buffer time</td> | |
| **15:55~16:00** | <td colspan="2" align="center">**Closing Remarks**</td> | |

# Data-Driven Designation and Implementation of Automated Guided Vehicles (ONSITE)

| Time | Title | Presenter/Author |
|---|---|---|
| WorkshopChairs: Jerry Chun-Wei Lin <jerrylin@ieee.org> | | |
| 10:30-10:50 | SP09201: An Efficient and Secured Energy Management System for Automated Guided Vehicles | Jia-Hao Syu, Jerry ChunWei Lin, and Dariusz Mrozek |
| 10:50-11:10 | SP09202: Botnet Attack Intrusion Detection In IoT Enabled Automated Guided Vehicles | Sumaiya Shaikh, Ch. Rupa, Gautam Srivastava, and Thippa Reddy Gadekallu |
| 11:10-11:30 | SP09205: Double-Environmental Q-Learning for Energy Management System in Smart Grid | Jia-Hao Syu, Jerry ChunWei Lin, and Philip S. Yu |
| 11:30-11:50 | SP09207: General Concepts of a Simulation Method for Automated Guided Vehicle in Industry 4.0 | Dariusz Marek, Piotr Biernacki, Jakub Szygula, and Adam Domanski |
| 11:50-12:10 | SP09208: Analysis of web-based geo-visualization methods applied for Automated Guided Vehicle using Satellite Navigation Systems | Jakub Szygula, Dariusz Marek, Adam Domanski, Lukasz Sobczak, and Piotr Pawlas |
| 12:10-12:30 | SP09209: Adaptive Data Forwarding for a Platoon based Vehicular Named Data Networking | Anu Kaushik, Deepanshu Garg, Anushka Nehra, Mohamed Baza, and Gautam Srivastava |
| Lunch Break | | |
| 14:00-14:20 | SP09210: Real-time testing of vision-based systems for AGVs with ArUco markers | Katarzyna Filus, Lukasz Sobczak, Joanna Domanska, Adam Domanski, and Rafal Cupek |
| 14:20-14:40 | SP09212: Communication between AGV and standalone station via EtherCAT using WiFi - proof of concept | Ireneusz Smolka, Jacek Stoj, Piotr Gaj, and Marcin Fojcik |
| 14:40-15:00 | SP09213: Neuro-heuristic pallet detection for AGV navigation | Katarzyna Prokop, Dawid Polap, and Gautam Srivastava |
| 15:00-15:20 | SP09215: Automated Guided Vehicles Challenges for Artificial Intelligence | Rafal Cupek, Jerry ChunWei Lin, and J. H. Syu |
| 15:20-15:40 | SP09216: Stream data clustering for engineering applications a use case of autonomous guided vehicles | Tomasz Steclik, Rafal Cupek, and Marek Drewniak |
| 15:40-16:00 | SP09217: The forecast of the AGV battery discharging via the machine learning methods | Olena Pavliuk, Tomasz Steclik, and Piotr Biernacki |
| Closing Remarks | | |

# Big Data Cup Challenges

| | Monday 19th of Dec. 2022<br>GAN-Based Training for Binary Classifier<br>*WorkshopChairs: Dr. Hussam Al Hamadi* | |
|---|---|---|
| **Time (+9 GMT)** | **Title** | **Presenter/Author** |
| 2:00 PM – 2:20 PM | **SC04201**: *Performance Enhancement of Malware Classifiers Using Generative Adversarial Networks* | Sunghyon Kyeong |
| 2:20 PM – 2:40 PM | **SC04205**: *OOG- Optuna Optimized GAN Sampling Technique for Tabular Imbalanced Malware Data* | S.M Towhidul Islam Tonmoy |
| 2:40 PM – 3:00 PM | **SC04203**: *Using Generative Adversarial Networks to Simulate System Calls of Malicious Android Processes* | Hamad Alsheraifi |
| 3:00 PM – 3:20 PM | **SC04202**: *A Visualized Malware Detection Framework with CNN and Conditional GAN* | Fang Wang |
| 3:20 PM – 3:40 PM | **SC04204**: *Using GAN to Generate Malicious Samples Suitable for Binary Classifier Training* | Sławomir Pioroński |

# Posters

| Poster ID | Accept Posters |
|---|---|
| P202 | Aparna Varde, Jianyu Liang, Zhaotong Yang, and Richard Sisson, *Ishikawa, JESS, and Visual Analytics for Engineering*<br>**Author Email(s):** vardea@montclair.edu, jianyul@wpi.edu, zyang6@wpi.edu, sisson@wpi.edu<br>**Contact Person:** Aparna Varde <vardea@montclair.edu> |
| P203 | Youjin Lee and Yonghan Roh, *Regression Yield Analysis for Semiconductor Manufacturing Based on Fabrication Conditions*<br>**Author Email(s):** youjinleee@g.skku.edu, yhroh@skku.edu<br>**Contact Person:** Youjin Lee <youjinleee@g.skku.edu> |
| P206 | Vincent Boudart and Maxime Fays, *ALBUS : A machine learning algorithm for gravitational wave burst searches*<br>**Author Email(s):** vboudart@uliege.be, Maxime.Fays@uliege.be<br>**Contact Person:** Vincent Boudart <vboudart@uliege.be> |
| P207 | Eric Bax, Arundhyoti Sarkar, and Natalie Bax, *Multiple Statistics: Marking Significance*<br>**Author Email(s):** ebax@yahooinc.com, arundhyoti@yahooinc.com, baxn@carleton.edu<br>**Contact Person:** Eric Bax <ebax@yahooinc.com> |
| P208 | Abdulkabir Abdulraheem and Im Jung, *A Resident Recognition Using Real and Thermal Image Pairs*<br>**Author Email(s):** iyjung@ee.knu.ac.kr, pooba0@daum.net<br>**Contact Person:** Im Jung <iyjung@ee.knu.ac.kr> |
| P209 | Eric Bax and Charlotte Bax, *Even Large Samples Have Challenges: Rare Events, Gaps, and Tails*<br>**Author Email(s):** ebax@yahooinc.com, 23baxc@flintridgeprep.org<br>**Contact Person:** Eric Bax <ebax@yahooinc.com> |

| P211 | Raid Lafta, Hisham Alshaheen, Biao Wang, Xiaohui Tao, Lingling Li, Kexin Zhang, and Ji Zhang, *Fast Fourier Transform and Ensemble Model to Classify Epileptic EEG Signals*<br>**Author Email(s):** Raidluaibi.lafta@utq.edu.iq, h.alshaheen@utq.edu.iq, Xiaohui.tao@usq.edu.au, lilingling@zua.edu.cn, kxzhang@zua.edu.cn, Ji.zhang@usq.edu.au<br>**Contact Person:** Raid Lafta <Raidluaibi.lafta@utq.edu.iq> |
|---|---|
| P212 | Sunhee Kim, Young-suk Lee, and Chang-Yong Lee, *Computational method of database construction for genetic variant calling*<br>**Author Email(s):** clee@kongju.ac.kr, king@kongju.ac.kr, youngl@kaist.ac.kr<br>**Contact Person:** Chang-Yong Lee <clee@kongju.ac.kr> |
| P213 | Giovanni Ciaramella, Giacomo Iadarola, Fabio Martinelli, Francesco Mercaldo, and Antonella Santone, *Continuous and Silent User Authentication Through Mouse Dynamics and Explainable Deep Learning: A Proposal*<br>**Author Email(s):** g.ciaramella1@studenti.unimol.it, giacomo.iadarola@iit.cnr.it, fabio.martinelli@iit.cnr.it, francesco.mercaldo@iit.cnr.it, antonella.santone@unimol.it<br>**Contact Person:** Giovanni Ciaramella <g.ciaramella1@studenti.unimol.it> |
| P216 | Seungshik Kang, Sangmo Lee, and Minghaeng Lee, *Word Phrase Analysis of High-frequency Ngrams in the Large-scale Korean Law Corpus*<br>**Author Email(s):** sskang@kookmin.ac.kr, peacekeeper@klri.re.kr, leemh@yonsei.ac.kr<br>**Contact Person:** Seungshik Kang <sskang@kookmin.ac.kr> |
| P217 | Ning Fang and P. Srinivasa Pai, *A Statistical Correlation Analysis of Big Data in Machining*<br>**Author Email(s):** ning.fang@usu.edu<br>**Contact Person:** Ning Fang <ning.fang@usu.edu> |
| P219 | Ansgar Lößer, Joel Witzke, Florian Schintke, and Björn Scheuermann, *BottleMod: Modeling Data Flows and Tasks for Fast Bottleneck Analysis*<br>**Author Email(s):** witzke@zib.de, ansgar.loesser@kom.tu-darmstadt.de, schintke@zib.de, scheuermann@kom.tu-darmstadt.de<br>**Contact Person:** Joel Witzke <witzke@zib.de> |
| P220 | Su-yeon Doo and Kee-Won Kwon, *Dynamic Power Reduction of TCAM Using Selective Precharging of Match Lines*<br>**Author Email(s):** dsy613@g.skku.edu, keewkwon@skku.edu<br>**Contact Person:** Su-yeon Doo <dsy613@g.skku.edu> |
| P221 | Jieling Wu and Mitsugu Saito, *Study on Statistical Analysis of Post-Earthquake Road Use Recovery and Its Application*<br>**Author Email(s):** wujl0612@yahoo.co.jp, mitsugu@iwate-u.ac.jp<br>**Contact Person:** Jieling Wu <wujl0612@yahoo.co.jp> |
| P222 | Dongyeon Kim, Yejin Kan, and Gangman Yi, *Novel Patrol Route Optimization Method Based on Bigdata Analysis*<br>**Author Email(s):** gangman@dongguk.edu, dongyeon@mme.dongguk.edu, yejin@mme.dongguk.edu<br>**Contact Person:** Gangman Yi <gangman@dongguk.edu> |
| P223 | SOJIN YOON, SOONMIN LEE, and YURA KWON, *Prediction of Postnatal Growth Failure in Very Low Birth Weight Infants using Artificial Intelligence Model*<br>**Author Email(s):** smlee@yuhs.ac, sojinyoon@yuhs.ac, yura970720@gmail.com<br>**Contact Person:** SOONMIN LEE <smlee@yuhs.ac> |
| P224 | Hyejin Kim, Dongseong Kim, Jin-Hee Cho, Terrence Moore, Frederica Nelson, and Hyuk Lim, *Continual Learning with Network Intrusion Dataset*<br>**Author Email(s):** hlim@kentech.ac.kr, hyejinkim@gist.ac.kr, dan.kim@uq.edu.au, jicho@vt.edu, terrence.j.moore.civ@army.mil, frederica.f.nelson.civ@army.mil<br>**Contact Person:** Hyuk Lim <hlim@kentech.ac.kr> |
| P225 | Irina Chernyavsky, Aparna Varde, and Simon Razniewski, *CSK-Detector: Commonsense in object detection*<br>**Author Email(s):** vardea@montclair.edu, chernyavskyi1@montclair.edu, srazniew@mpi-inf.mpg.de<br>**Contact Person:** Aparna Varde <vardea@montclair.edu> |
| P226 | Sangkeun Lee, Supriya Chinthavali, Sarah Tennille, Junghoon Chae, Anika Tabassum, Varisara Tansakul, Daniel Redmon, Robert Moncrief, and Aaron Myers, *Graph-based Cascading Impact Estimation for Identifying Crucial Infrastructure Components*<br>**Author Email(s):** lees4@ornl.gov<br>**Contact Person:** Sangkeun Lee <lees4@ornl.gov> |

| P228 | Wei Quan, Zhijian Li, and Wonkyung Kim, *Implicit Discriminatory Bias Reflected in Yelp Reviews - A case study*<br>**Author Email(s):** weiquan@uic.edu.cn, zhijianli@uic.edu.cn, wonkyungkim@uic.edu.cn<br>**Contact Person:** Wei Quan <weiquan@uic.edu.cn> |
|------|---|
| P229 | Zhongliang Zhou, Nathaniel Hitt, Benjamin Letcher, Weili Shi, and Sheng Li, *Pigmentation-based Visual Learning for Salvelinus fontinalis Individual Re-identification*<br>**Author Email(s):** shengli@virginia.edu, zz42551@uga.edu, nhitt@usgs.gov, bletcher@usgs.gov, rhs2rr@virginia.edu<br>**Contact Person:** Sheng Li <shengli@virginia.edu> |
| P230 | Taewhi Lee, Kihyuk Nam, Choon Seo Park, and Sung-Soo Kim, *Exploiting Machine Learning Models for Approximate Query Processing*<br>**Author Email(s):** taewhi@etri.re.kr, nam@etri.re.kr, parkcs@etri.re.kr, sungsoo@etri.re.kr<br>**Contact Person:** Taewhi Lee <taewhi@etri.re.kr> |
| P231 | Sohee Ban, Hyo Suk Nam, and Eunjeong Park, *Detecting Paralysis of Stroke Symptom in Video: Transfer Learning with Gated Recurrent Unit using Public Big Data of Facial Images*<br>**Author Email(s):** eunjeong.ej@gmail.com, bsh1004664@yuhs.ac, hsnam@yuhs.ac<br>**Contact Person:** Eunjeong Park <eunjeong.ej@gmail.com> |
| P234 | Aviv Yehezkel and Eyal Elyashiv, *A GNN-based Approach for Detecting Network Anomalies from Small Traffic Samples*<br>**Author Email(s):** aviv@cynamics.ai<br>**Contact Person:** Aviv Yehezkel <aviv@cynamics.ai> |
| P235 | Daniel Cheriyan, Jurng-Jae Yee, and jae-ho Choi, *CONSTRUCTION WORKER HEALTH AND SAFETY MONITORING SYSTEM USING BRAINWAVE AND POLLUTION DATA*<br>**Author Email(s):** jaehochoi@dau.ac.kr, daniel@donga.ac.kr, jjyee@dau.ac.kr<br>**Contact Person:** jae-ho Choi <jaehochoi@dau.ac.kr> |
| P237 | MD Hasan, Hoda Maleki, and Gagan Agrawal, *Securing Pseudonym Schemes for Vehicular Privacy*<br>**Author Email(s):** hmaleki@augusta.edu, mdhasan@augusta.edu, gagrawal@augusta.edu<br>**Contact Person:** Hoda Maleki <hmaleki@augusta.edu> |
| P238 | Ramses J. Sotelo Jimenez, Bhavya Medishetty, and Byron J. Gao, *COMPA: A Comparative Retrieval and Analytical Engine for Consumer Products*<br>**Author Email(s):** bgao@txstate.edu, rjs144@txstate.edu, b_m156@txstate.edu<br>**Contact Person:** Byron J. Gao <bgao@txstate.edu> |
| P240 | Koji Iwanuma and Ryo Hinata, *A Fast On-Line ε-Approximation Algorithm for Mining Strongly Closed Itemsets*<br>**Author Email(s):** iwanuma@yamanashi.ac.jp<br>**Contact Person:** Koji Iwanuma <iwanuma@yamanashi.ac.jp> |
| P241 | YOUNG JIN KIM, CHUNHYUN PAIK, and YONGJOO CHUNG, *Outlier Elimination Procedure for Power Curve Modeling of Wind Generation*<br>**Author Email(s):** youngk@pknu.ac.kr, chpaik@deu.ac.kr, chungyj@bufs.ac.kr<br>**Contact Person:** YOUNG JIN KIM <youngk@pknu.ac.kr> |
| P242 | Keun-Chang Kwak and Yeong-Hyeon Byeon, *Performance Comparison for Deep Learning-based Personal Identification from the database of ECG and EMG signals*<br>**Author Email(s):** kwak@chosun.ac.kr<br>**Contact Person:** Keun-Chang Kwak <kwak@chosun.ac.kr> |
| P243 | Kangmin Kim and Sangwhan Cha, *Soccer Ball Flow Analysis based on Markov Chain and PCA*<br>**Author Email(s):** kim785440@sas.edu.sg, scha@harrisburgu.edu<br>**Contact Person:** Kangmin Kim <kim785440@sas.edu.sg> |
| P244 | Ting Jiang, Ting Yu, Zexian Hong, Zujie Ren, and Ji Zhang, *A Parallel Framework for Streaming Graphs Computing*<br>**Author Email(s):** jiangting@zhejianglab.com, yuting@zhejianglab.com, HONG0223@e.ntu.edu.sg, zujie.rzj@zhejianglab.com, zhangji77@gmail.com<br>**Contact Person:** Ting Jiang <jiangting@zhejianglab.com> |
| P245 | Vo Khanh Ngan Tran, Teruhisa Hochin, Hiroki Nomiya, Hideya Nakanishi, and Mamoru Shoji, *Prediction of Plasma Discharge Video by Convolutional Long Short-Term Memory*<br>**Author Email(s):** d1821503@edu.kit.ac.jp, hochin@kit.ac.jp, nomiya@kit.ac.jp, |

| | |
|---|---|
| | nakanishi.hideya@nifs.ac.jp, shohji.mamoru@nifs.ac.jp<br>**Contact Person:** Vo Khanh Ngan Tran <d1821503@edu.kit.ac.jp> |
| P246 | Toshiyuki Shimono, *Geographic Reason by Hourly Population of 500-Meter-Grid Statistics over Japan Why Tokyo Failed to Eradicate COVID-19 in Mid 2020*<br>**Author Email(s):** tulamili@gmail.com<br>**Contact Person:** Toshiyuki Shimono <tulamili@gmail.com> |
| P247 | Yiqun Jin and Siya Bao, *ML-Based Trading Strategy for Short-Term Price Reactions on Earnings Announcement Reports*<br>**Author Email(s):** jinguanyan@asagi.waseda.jp, siya.bao@togawa.cs.waseda.ac.jp<br>**Contact Person:** Yiqun Jin <jinguanyan@asagi.waseda.jp> |
| P248 | Anika Tabassum, Nils Stenvig, Sangkeun Lee, and Supriya Chinthavali, *Rule-based Quantification to Identify Crucial Power System Components for Mitigating Disaster Impact*<br>**Author Email(s):** tabassuma@ornl.gov, stenvignm@ornl.gov, lees4@ornl.gov, chinthavalis@ornl.gov<br>**Contact Person:** Anika Tabassum <tabassuma@ornl.gov> |
| P249 | Anastasios Alexandridis, Ghassan Al-Sumaidaee, Rami Alkhudary, and Zeljko Zilic, *A Stylized Presence Detection System in the Era of Blockchain and Big Data*<br>**Author Email(s):** anastasios.alexandridis@mail.mcgill.ca, ghassan.al-sumaidaee@mail.mcgill.ca, rami.alkhudary@u-paris2.fr, zeljko.zilic@mcgill.ca<br>**Contact Person:** Anastasios Alexandridis <anastasios.alexandridis@mail.mcgill.ca> |
| P250 | Renhe Jiang, Zekun Cai, Zhaonan Wang, Chuang Yang, Zipei Fan, Quanjun Chen, Kota Tsubouchi, Xuan Song, and Ryosuke Shibasaki, *Yahoo! Bousai Crowd Data: A Large-Scale Crowd Density and Flow Dataset in Tokyo and Osaka*<br>**Author Email(s):** jiangrh@csis.u-tokyo.ac.jp, caizekun@csis.u-tokyo.ac.jp, znwang@csis.u-tokyo.ac.jp, chuang.yang@csis.u-tokyo.ac.jp, fanzipei@iis.u-tokyo.ac.jp, chen1990@csis.u-tokyo.ac.jp, ktsubouc@yahoo-corp.jp, songxuan@csis.u-tokyo.ac.jp, shiba@csis.u-tokyo.ac.jp<br>**Contact Person:** Renhe Jiang <jiangrh@csis.u-tokyo.ac.jp> |
| P251 | Heehun Jeong, Giup Seo, and Euiseok Hwang, *Lossless Data Compression with Bit-back Coding on Massive Smart Meter Data*<br>**Author Email(s):** desperadohh@gm.gist.ac.kr, guseo1120@gm.gist.ac.kr, euiseokh@gist.ac.kr<br>**Contact Person:** Heehun Jeong <desperadohh@gm.gist.ac.kr> |
| P252 | Hyung-Jin Kim, Kwang-Sik Kim, Min-Seok Kim, and Jang-Hyun Lee, *Development of Fault Detection Algorithm using LP Pump Failure Data*<br>**Author Email(s):** gudwls4369@naver.com, kskim00@gmail.com, rhiho1008@nate.com, jh_lee@inha.ac.kr<br>**Contact Person:** Hyung-Jin Kim <gudwls4369@naver.com> |
| P253 | Zi Hen Lin, Ziwei Wang, Minzhu Zhao, Yunya Song, and Liang Lan, *An AI-based System to Assist Human Fact-Checkers for Labeling Cantonese Fake News on Social Media*<br>**Author Email(s):** lanliang@hkbu.edu.hk, 17211956@life.hkbu.edu.hk, 21482837@life.hkbu.edu.hk, uczcmz0@ucl.ac.uk, yunyasong@hkbu.edu.hk<br>**Contact Person:** Liang Lan <lanliang@hkbu.edu.hk> |
| P254 | Guanrong Li, Ziwei Wang, Minzhu Zhao, Yunya Song, and Liang Lan, *Sentiment Analysis of Political Posts on Hong Kong Local Forums Using Fine-Tuned mBERT*<br>**Author Email(s):** lanliang@hkbu.edu.hk, grli@smail.nju.edu.cn, 21482837@life.hkbu.edu.hk, uczcmz0@ucl.ac.uk, yunyasong@hkbu.edu.hk<br>**Contact Person:** Liang Lan <lanliang@hkbu.edu.hk> |
| P256 | Oyun Kwon, Junhwan Kwon, Kyeong Teak Oh, and Sun K. Yoo, *Interoperable Reference Model for Public Data in Medical Imaging*<br>**Author Email(s):** sunkyoo@yuhs.ac, koy9303@yuhs.ac, kjh0442@yuhs.ac, okt2704@yuhs.ac<br>**Contact Person:** Sun K. Yoo <sunkyoo@yuhs.ac> |
| P257 | Junhwan Kwon, Jiwon choi, Oyun Kwon, Kyeong Teak Oh, Hayoung Kim, Jaesuk Kim, and Sun K. Yoo, *Breathing detection using thermal facial landmark with DICOM file*<br>**Author Email(s):** sunkyoo@yuhs.ac, kjh0442@yuhs.ac, zx7398@yuhs.ac, koy9303@yuhs.ac, okt2704@yuhs.ac, hykim9359@yuhs.ac, rlawotjr1207@yuhs.ac<br>**Contact Person:** Sun K. Yoo <sunkyoo@yuhs.ac> |
| P258 | JongSeok Jeon, SeungHo Kim, and JooHee Oh, *How Do You watch Music Video : Watching or Listening?*<br>**Author Email(s):** jongseok@handong.ac.kr, kshkjh923@gmail.com, jooheeoh@handong.edu<br>**Contact Person:** JongSeok Jeon <jongseok@handong.ac.kr> |

| | |
|---|---|
| P259 | Masahiro Nakahara, Yuichi Mashiba, Ryusuke Miyamoto, Yuki Fujita, Hisashi Ishida, and Keiichi Zempo, *Inspection of unexpected defective products by semi-supervised learning based on a probability density function in high-yield food production*<br>**Author Email(s):** m.nakahara@ishida-tec.co.jp, zempo@iit.tsukuba.ac.jp<br>**Contact Person:** Masahiro Nakahara <m.nakahara@ishida-tec.co.jp> |
| P260 | Kangmin Kim and Sangwhan Cha, *Towards Developing Face Analysis System based PCA*<br>**Author Email(s):** kim785440@sas.edu.sg, scha@harrisburgu.edu<br>**Contact Person:** Kangmin Kim <kim785440@sas.edu.sg> |
| P261 | Pedro Augusto Da S. E S. Miranda and Plinio P. Morita, *A Proposed Personal Data Sovereignty Inter-Organizational Governance Framework for Public Health Research*<br>**Author Email(s):** pmiranda@uwaterloo.ca, plinio.morita@uwaterloo.ca<br>**Contact Person:** Pedro Augusto Da S. E S. Miranda <pmiranda@uwaterloo.ca> |
| P262 | Janusz Bobulski and Mariusz Kubanek, *A method of cleaning data from IoT devices in Big data systems*<br>**Author Email(s):** januszb@icis.pcz.pl, mkubanek@icis.pcz.pl<br>**Contact Person:** Janusz Bobulski <januszb@icis.pcz.pl> |
| P263 | Akinori Asahara, Yoshihiro Osakabe, Karin Tsuda, and Hidekazu Morita, *Interactive Support System for Idea Divergence-Convergence Iteration*<br>**Author Email(s):** akinori.asahara.bq@hitachi.com, yoshihiro.osakabe.fj@hitachi.com, karin.tsuda.sc@hitachi.com, hidekazu.morita.tq@hitachi.com<br>**Contact Person:** Akinori Asahara <akinori.asahara.bq@hitachi.com> |
| P264 | Ezekiel Young and Hady Lauw, *Mining Competitively-Priced Bundle Configurations*<br>**Author Email(s):** eoyoung.2019@smu.edu.sg, hadywlauw@smu.edu.sg<br>**Contact Person:** Ezekiel Young <eoyoung.2019@smu.edu.sg> |
| P267 | M M Mahabubur Rahman and Jelena Tesic, *Evaluating Hybrid Approximate Nearest Neighbor Indexing and Search (HANNIS) for High-dimensional Image Feature Search*<br>**Author Email(s):** toufik@txstate.edu, jtesic@txstate.edu<br>**Contact Person:** M M Mahabubur Rahman <toufik@txstate.edu> |
| P268 | Shahan Salim, Irfhana Zakir Hussain, Jasleen Kaur, and Plinio Morita, *Air Pollution Surveillance Ecosystem: A Big Data Approach to Monitoring Adverse Health Outcomes for Public Health Interventions*<br>**Author Email(s):** shahan.salim@uwaterloo.ca, irfhana.zakirhussain@uwaterloo.ca, jasleen.kaur@uwaterloo.ca, plinio.morita@uwaterloo.ca<br>**Contact Person:** Shahan Salim <shahan.salim@uwaterloo.ca> |
| P269 | Minseok Jeong, Giup Seo, and Euiseok Hwang, *Age of Information Optimization by Deep Reinforcement Learning for Random Access in Machine Type Communication*<br>**Author Email(s):** dsa950115@gm.gist.ac.kr, guseo1120@gm.gist.ac.kr, euiseokh@gist.ac.kr<br>**Contact Person:** Minseok Jeong <dsa950115@gm.gist.ac.kr> |
| P270 | Yeongmin Kim, Yoon Jung Ko, and Seung Jin Oh, *Evaluation and validation of an artificial neural network for predicting the performance of a desiccant coated heat exchanger*<br>**Author Email(s):** ohs8680@kitech.re.kr, ymkin5321@kitech.re.kr, koyj486@kitech.re.kr<br>**Contact Person:** Seung Jin Oh <ohs8680@kitech.re.kr> |
| P271 | Yang Zhang, Zhen Wang, Ting Yu, Mingming Lu, Zujie Ren, and Ji Zhang, *Knowledge Tracing Based on Gated Heterogeneous Graph Convolutional Networks*<br>**Author Email(s):** zhangyang_zy@zhejianglab.com, wangzhen@zhejianglab.com, yuting@zhejianglab.com, mingminglu@csu.edu.cn, zujie.rzj@zhejianglab.com, zhangji77@gmail.com<br>**Contact Person:** Yang Zhang <zhangyang_zy@zhejianglab.com> |
| P272 | Jacopo Fior, Thomas Favale, Luca Cagliero, Danilo Giordano, Marco Mellia, Elena Baralis, Silvia Ronchiadin, Paolo Baracco, and Dario Moncalvo, *Legal Entity Disambiguation For Financial Crime Detection*<br>**Author Email(s):** jacopo.fior@polito.it, thomas.favale@polito.it, luca.cagliero@polito.it, danilo.giordano@polito.it, marco.mellia@polito.it, elena.baralis@polito.it, silvia.ronchiadin@intesasanpaolo.com, paolo.baracco@intesasanpaolo.com, dario.moncalvo@intesasanpaolo.com<br>**Contact Person:** Jacopo Fior <jacopo.fior@polito.it> |
| P273 | Mingchen Lu, Chen Tang, and Peiquan Jin, *Revisiting LSM-Tree-Based Key-Value Stores for ZNS SSDs*<br>**Author Email(s):** jpq@ustc.edu.cn, lmc123@mail.ustc.edu.cn, tcinustc@mail.ustc.edu.cn<br>**Contact Person:** Peiquan Jin <jpq@ustc.edu.cn> |

| P274 | Kyeongsu Byun, Joonho Kwon, and Goo Kim, *Fairness improvement technology and visualization services for binary classification datasets collected on a batch basis*<br>**Author Email(s):** sxdcfv3955@gmail.com, jhkwon@pusan.ac.kr, daekeun7@gmail.com<br>**Contact Person:** Kyeongsu Byun <sxdcfv3955@gmail.com> |
|---|---|
| P275 | Chen Tang, Can Wang, Mingchen Lu, and Peiquan Jin, *An Efficient In-Memory Analytics System Based on Persistent Memory*<br>**Author Email(s):** jpq@ustc.edu.cn, tcinustc@mail.ustc.edu.cn, wvolcano@mail.ustc.edu.cn, lmc123@mail.ustc.edu.cn<br>**Contact Person:** Peiquan Jin <jpq@ustc.edu.cn> |
| P276 | Allison Hsu and Ying-Feng Hsu, *Comparison of Teenagers' Writing Using Word Clouds and Analyzing Engines*<br>**Author Email(s):** allison.lc.hsu@gmail.com, yf.hsu@cmc.osaka-u.ac.jp<br>**Contact Person:** Allison Hsu <allison.lc.hsu@gmail.com> |
| P277 | Zhen Wang, Madison Cooley, Yang Zhang, Chao Lan, and Ji Zhang, *Debiased Learning of Self-Labeled Twitter Data for User Demographic Prediction*<br>**Author Email(s):** wangzhen@zhejianglab.com, mcooley@uwyo.edu, zhangyang_zy@zhejianglab.com, clan@uwyo.edu, zhangji77@gmail.com<br>**Contact Person:** Zhen Wang <wangzhen@zhejianglab.com> |
| P278 | Hyogyun An, Huiju Kim, Hyuntae Park, and Victoire Cyrot, *Perceptual-IQ: Visual Commonsense Reasoning about Perceptual Imagination*<br>**Author Email(s):** yk6525@korea.ac.kr, haena0320@korea.ac.kr, pht0639@korea.ac.kr, 2021021585@korea.ac.kr<br>**Contact Person:** Hyogyun An <yk6525@korea.ac.kr> |
| P279 | Zenghui Xu, Ting Yu, Xingyun Hong, Mingzhang Li, Yang Zhang, Zujie Ren, and Ji Zhang, *IDGMS: a One-Stop Graph Mining System for Infectious Diseases*<br>**Author Email(s):** xuzenghui@zhejianglab.com, yuting@zhejianglab.com, xyhong@zhejianglab.com, 1244437539@qq.com, zhangyang_zy0823@163.com, zujie.rzj@zhejianglab.com, Ji.Zhang@usq.edu.au<br>**Contact Person:** Zenghui Xu <xuzenghui@zhejianglab.com> |
| P280 | Dohee Kim and Hyerim Bae, *TDTA: A New Hybrid Framework for Long-term Forecasting of Container Volume*<br>**Author Email(s):** hrbae@pusan.ac.kr, kimdohee@pusan.ac.kr<br>**Contact Person:** Hyerim Bae <hrbae@pusan.ac.kr> |
| P283 | Emilija Strelcenia and Dr Simant Prakoonwit, *GAN-BASED DATA AUGMENTATION FOR CREDIT CARD FRAUD DETECTION*<br>**Author Email(s):** s5223247@bournemouth.ac.uk, strelceniae@bournemouth.ac.uk<br>**Contact Person:** Emilija Strelcenia <s5223247@bournemouth.ac.uk> |
| P284 | Soumaya El Herry, Naoya Tanaka, and Akihiro Fujii, *Marine Water Quality Analysis of Kesennuma Bay in the Aftermath of the 2011 Great East Japan Earthquake and Tsunami*<br>**Author Email(s):** soumaya.herry@gmail.com, decarch19@gmail.com, fujii@hosei.ac.jp<br>**Contact Person:** Soumaya El Herry <soumaya.herry@gmail.com> |
| P285 | Kevin I-Kai Wang, Xiaozhou Ye, and Kouichi Sakurai, *Federated Learning with Clustering-Based Participant Selection for IoT Applications*<br>**Author Email(s):** kevin.wang@auckland.ac.nz, xye685@aucklanduni.ac.nz, sakurai@inf.kyushu-u.ac.jp<br>**Contact Person:** Kevin I-Kai Wang <kevin.wang@auckland.ac.nz> |
| P286 | jae-ho Choi, Jin Young Choi, and Eun-Ju Ha, *Data Mining Approach to determine injured Area of the Construction Worker using Accident Loss Cost*<br>**Author Email(s):** jaehochoi@dau.ac.kr, jinyoung0725@gmail.com, 19731jmj@donga.ac.kr<br>**Contact Person:** jae-ho Choi <jaehochoi@dau.ac.kr> |
| P288 | Mikyung Je, Myeongji Cho, Yoon Jung Kim, and Hyeon S. Son, *Exploring the mutation tendency of a specific region in the virus taxonomic groups using data mining techniques*<br>**Author Email(s):** hss2003@snu.ac.kr, jemi57@snu.ac.kr, yummy0112@naver.com, okilinko@snu.ac.kr<br>**Contact Person:** Hyeon S. Son <hss2003@snu.ac.kr> |
| P289 | Seungnam Han, Haewon Lee, Seungwook Yoon, and Euiseok Hwang, *Localization of IoT Devices based on Spatially Separated Multi-Antenna Radio Frequency Channel State Information*<br>**Author Email(s):** euiseokh@gist.ac.kr, snhan0911@gm.gist.ac.kr, haewonlii@gm.gist.ac.kr, ysw1207@gist.ac.kr<br>**Contact Person:** Euiseok Hwang <euiseokh@gist.ac.kr> |

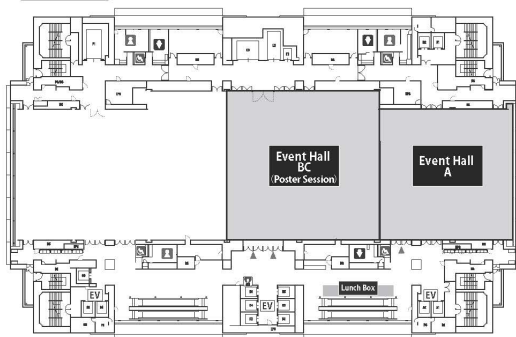| | |
|---|---|
| P290 | Seonhyeong Kim and Young-Woo Kwon, *Construction of Disaster Knowledge Graphs to Enhance Disaster Resilience*<br>**Author Email(s):** kimsh9510@gmail.com, ywkwon@knu.ac.kr<br>**Contact Person:** Seonhyeong Kim <kimsh9510@gmail.com> |
| P291 | Eunbi Cho, Sungil Hong, Hyeseo Yoon, Eunsung Cho, Jinho Shim, JooHee Oh, and Junghyun Kim, *Data-driven Approach using Unsupervised Learning for Detecting Anomalies in Facility Operations*<br>**Author Email(s):** junghyun.kim@handong.edu, 22000724@handong.ac.kr, jooheeoh@handong.ac.kr<br>**Contact Person:** Junghyun Kim <junghyun.kim@handong.edu> |
| P292 | Irshad Khan and Young-Woo Kwon, *Attention-based Malware Detection of Android Applications*<br>**Author Email(s):** irshad.cs@knu.ac.kr, ywkwon@knu.ac.kr<br>**Contact Person:** Irshad Khan <irshad.cs@knu.ac.kr> |
| P293 | Jangsoo Lee, Jae-Heon Sim, Jae-Kwang Ahn, and Young-Woo Kwon, *An Open Dataset for Deep Learning-based Earthquake Detection using MEMS Sensors*<br>**Author Email(s):** ywkwon@knu.ac.kr, dellhart@knu.ac.kr, zh025700@knu.ac.kr, propjk@korea.kr<br>**Contact Person:** Young-Woo Kwon <ywkwon@knu.ac.kr> |
| P294 | Kihyuk Nam, Taewhi Lee, Sung-Soo Kim, Choon Seo Park, Taek Yong Nam, and Insik Shin, *An Efficient Data Analysis for Edge-Enabled Distributed Environments using Tractable Probabilistic Models*<br>**Author Email(s):** nam@etri.re.kr, taewhi@etri.re.kr, sungsoo@etri.re.kr, parkcs@etri.re.kr, tynam@etri.re.kr, ishin@kaist.ac.kr<br>**Contact Person:** Kihyuk Nam <nam@etri.re.kr> |
| P295 | Yujin Choi and JooHee Oh, *Does ESG Score and Intangible Asset improve Financial performance?: Machine Learning Model of KOSPI Enterprises*<br>**Author Email(s):** 21900754@handong.ac.kr, jooheeoh@handong.edu<br>**Contact Person:** Yujin Choi <21900754@handong.ac.kr> |
| P298 | jae-ho Choi, Eun-Ju Ha, and Jin Young Choi, *Data Mining Approach for Predicting the Fall Accident Cost in Construction Projects*<br>**Author Email(s):** jaehochoi@dau.ac.kr, 19731jmj@donga.ac.kr, jinyoung0725@gmail.com<br>**Contact Person:** jae-ho Choi <jaehochoi@dau.ac.kr> |
| P299 | Seongwoo Cho, Yujin Jung, Yongdeuk Seo, Kyuheon Jung, and Sungchul Choi, *Optimization Analysis and Standardization of Rolling Process with Machine Learning*<br>**Author Email(s):** sc82.choi@pknu.ac.kr, jsw6872@pukyong.ac.kr, wjd1dbwls@pukyong.ac.kr, sod7050@pukyong.ac.kr, kkyuhun94@pukyong.ac.kr<br>**Contact Person:** Sungchul Choi <sc82.choi@pknu.ac.kr> |
| P301 | HaEun Jang, SeungHo Kim, JongSeok Jeon, and JooHee Oh, *Visual Attributes of Thumbnails in Predicting Top YouTube Brand Channels: A Machine Learning Approach*<br>**Author Email(s):** jjangha@handong.ac.kr, kshkjh923@gmail.com, jongseok@handong.ac.kr, jooheeoh@handong.edu<br>**Contact Person:** HaEun Jang <jjangha@handong.ac.kr> |
| P302 | Xuyang Cao, Yang Cao, Masatoshi Yoshikawa, and Atsuyoshi Nakamura, *Boosting Utility of Differentially Private Streaming Data Release under Temporal Correlations*<br>**Author Email(s):** xuyang.cao.z1@elms.hokudai.ac.jp, yang@ist.hokudai.ac.jp, yoshikawa@i.kyoto-u.ac.jp, atsu@ist.hokudai.ac.jp<br>**Contact Person:** Xuyang Cao <xuyang.cao.z1@elms.hokudai.ac.jp> |
| P303 | Eunjeong Park, Hyo Suk Nam, and Jae-Woo Song, *Federated Learning Models using Flow Cytometry Data of Blood Test in Medical Decision Support*<br>**Author Email(s):** eunjeong.ej@gmail.com, hsnam@yuhs.ac<br>**Contact Person:** Eunjeong Park <eunjeong.ej@gmail.com> |
| P304 | Mohamed MAKHLOUF, *BIG: Big data Intelligence Governance framework*<br>**Author Email(s):** mohamed.makhlouf@essca.fr<br>**Contact Person:** Mohamed MAKHLOUF <mohamed.makhlouf@essca.fr> |
| P305 | DoHyun Kwon, SeongBin Ju, and JeongHun Kim, *Analysis of factors affecting the change of PM10 in Korea*<br>**Author Email(s):** 21700043@handong.ac.kr, 21700702@handong.ac.kr, 21800166@handong.ac.kr<br>**Contact Person:** DoHyun Kwon <21700043@handong.ac.kr> |
| P307 | Trong Nguyen, Youness Arjoune, Jonathan Schroeder, Dinesh Pillai, Stephen Teach, Shilpa Patel, and Raj Shekhar, *MACHINE LEARNING FOR AUTOMATED WHEEZE DETECTION IN CHILDREN*<br>**Author Email(s):** tnn@auscultechdx.com, yarjoune@childrensnational.org, |

| | |
|---|---|
| | jschroeder@childrensnational.org, DPillai@childrensnational.org, steach@childrensnational.org, SPatel@childrensnational.org, RShekhar@childrensnational.org<br>**Contact Person:** Trong Nguyen <tnn@auscultechdx.com> |
| P308 | Kota Kumakura, Masato Oguchi, Takeshi Kamiyama, and Saneyasu Yamaguchi, *CPU Usage Trends in Android Applications*<br>**Author Email(s):** cm22022@g.kogakuin.jp, oguchi@is.ocha.ac.jp, kami@nagasaki-u.ac.jp, sane@cc.kogakuin.ac.jp<br>**Contact Person:** Kota Kumakura <cm22022@g.kogakuin.jp> |
| P309 | Naoki Monji and Keita Ushida, *Making Karaoke Parties Lively by Reordering Songs Based on Pop Music Concert Program Data*<br>**Author Email(s):** ushida@cc.kogakuin.ac.jp, j016288@kute.tokyo<br>**Contact Person:** Keita Ushida <ushida@cc.kogakuin.ac.jp> |
| P310 | Ghassan Al-Sumaidaee, Rami Alkhudary, and Zeljko Zilic, *Decentralized Storage for Big Data in Healthcare between Reality and Ambition: IPFS and Sia*<br>**Author Email(s):** ghassan.al-sumaidaee@mail.mcgill.ca, rami.alkhudary@u-paris2.fr, zeljko.zilic@mcgill.ca<br>**Contact Person:** Ghassan Al-Sumaidaee <ghassan.al-sumaidaee@mail.mcgill.ca> |
| P311 | Sara Nieto Rodriguez, Bruno Altieri, and Pilar de Teodoro, *BIG DATA EXPLOITATION IN COSMOLOGY: THE EUCLID SCIENTIFIC ARCHIVE SYSTEM*<br>**Author Email(s):** sara.nieto@ext.esa.int, bruni.altieri@esa.int, pilar.teodoro@ext.esa.int<br>**Contact Person:** Sara Nieto Rodriguez <sara.nieto@ext.esa.int> |
| P312 | Tai-Yeon Ku, Wan-Ki Park, and Hoon Choi, *Energy Maestro–Transactive Energy Mechanism*<br>**Author Email(s):** kutai@etri.re.kr, kutai@etri.re.kr, kutai@etri.re.kr<br>**Contact Person:** Tai-Yeon Ku <kutai@etri.re.kr> |
| P313 | Quentin Petit, Nahid Emad, and Chong Li, *Distributed and parallel sparse computing for very large graph neural networks*<br>**Author Email(s):** petitquentin41@gmail.com, nahid.emad@uvsq.fr, ch.l@huawei.com<br>**Contact Person:** Quentin Petit <petitquentin41@gmail.com> |
| P315 | Chae-Won Lee, So-Hyun Cho, Jong-Hun Lee, Jun-Seong Lee, Sungchul Choi, and Won-Chul Seo, *Empirical Research in Anomaly Detection for Rotating Machines Diagnosis using Deep Learning*<br>**Author Email(s):** wcseo@pknu.ac.kr, oochaewonoo@gmail.com, sh.cho.020@gmail.com, whdgns4608@gmail.com, a01096982827@gmail.com, sc82.choi@pknu.ac.kr<br>**Contact Person:** Won-Chul Seo <wcseo@pknu.ac.kr> |
| P316 | Marc Cruz, Sherelene De Beleh, Russel Janairo, Yanbo Zhao, Isha Raheja, and Tingting Chen, *Mobility Scooter User Driving Behavior Classification based on Deep Neural Networks*<br>**Author Email(s):** tingtingchen@cpp.edu, marccruz@cpp.edu, sddebelen@cpp.edu, rsjanairo@cpp.edu, yanbozhao716@gmail.com<br>**Contact Person:** Tingting Chen <tingtingchen@cpp.edu> |
| P317 | Thilanka Munasinghe and Brandon Behlendorf, *Using Graph Neural Networks to Investigate the Relationship Between the Socioeconomic Factors and Emergency Medical Service (EMS) Median Response Time in New York City*<br>**Author Email(s):** munast@rpi.edu, bbehlendorf@albany.edu<br>**Contact Person:** Thilanka Munasinghe <munast@rpi.edu> |
| P318 | Junwoo Ha and Im Jung, *An autonomous maximum speed control considering boarding weight for safe e-scooter driving*<br>**Author Email(s):** iyjung@ee.knu.ac.kr, gkwnsdn0402@naver.com<br>**Contact Person:** Im Jung <iyjung@ee.knu.ac.kr> |
| P319 | Yuma Konno, Miyuki Nakano, and Masato Oguchi, *Efficient Data Selection Indicators for Updating Models under Data Drifted Environment*<br>**Author Email(s):** konno.yuma@is.ocha.ac.jp, miyuki@tsuda.ac.jp, oguchi@is.ocha.ac.jp<br>**Contact Person:** Yuma Konno <konno.yuma@is.ocha.ac.jp> |
| P322 | Daisuke Katayama, Kippei Kasai, and Takahiro Koita, *Migration Destination Selection Algorithm for Spot Instances using SPS*<br>**Author Email(s):** ctwh0127@mail4.doshisha.ac.jp, cgud1022@mail4.doshisha.ac.jp, tkoita@mail.doshisha.ac.jp<br>**Contact Person:** Daisuke Katayama <ctwh0127@mail4.doshisha.ac.jp> |

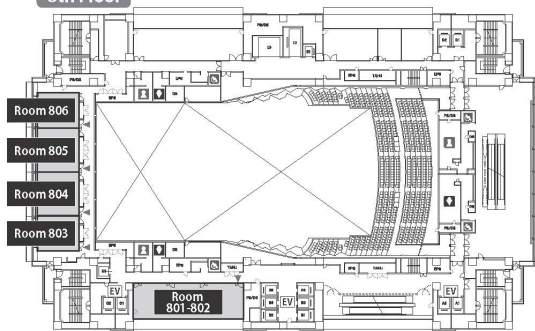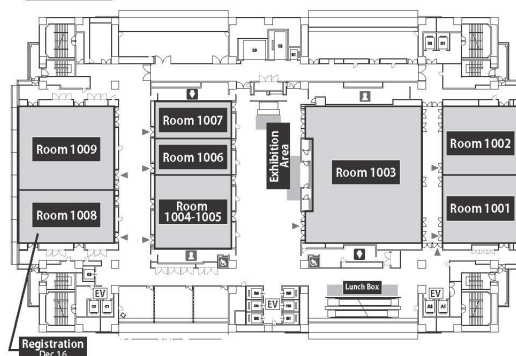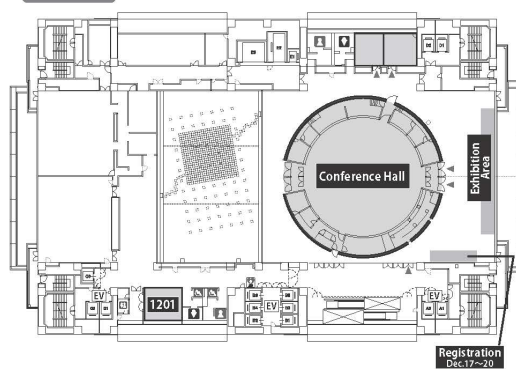| P323 | Dong-Hyun Kang, Jeong-Hwa Lee, Tae-Sun Yu, and Yushin Lee, *Data Driven Optimal Coil Batching for Annealing Lines with Eligibility Constraints* <br> **Author Email(s):** tsyu@pknu.ac.kr, xrgsnim@gmail.com, qpqp8558@gmail.com, ysl@pknu.ac.kr <br> **Contact Person:** Tae-Sun Yu <tsyu@pknu.ac.kr> |
|---|---|
| P324 | GeeHoon Lee and Im Jung, *A fire prevention by abnormal heat generation detection using LSTM* <br> **Author Email(s):** iyjung@ee.knu.ac.kr, wlgns4123@naver.com <br> **Contact Person:** Im Jung <iyjung@ee.knu.ac.kr> |

# Floor Layout

**Floor Layout**

**3rd Floor**

Event Hall BC (Poster Session)

Event Hall A

Lunch Box

EV  EV  EV

10

**8th Floor**

Room 806
Room 805
Room 804
Room 803

Room 801-802

EV  EV  EV

**10th Floor**

Room 1009
Room 1008

Room 1007
Room 1006
Room 1004-1005

Exhibition Area

Room 1003

Room 1002
Room 1001

Lunch Box

EV  EV  EV

Registration
Dec.16

**12th Floor**

Conference Hall

Exhibition Area

1201

EV  EV

Registration
Dec.17~20

11