

2023 IEEE 39th International Conference on Data Engineering Workshops (ICDEW 2023)

**Anaheim, California, USA
3 – 7 April 2023**



**IEEE Catalog Number: CFP2345A-POD
ISBN: 979-8-3503-2245-3**

**Copyright © 2023 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

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

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

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

IEEE Catalog Number:	CFP2345A-POD
ISBN (Print-On-Demand):	979-8-3503-2245-3
ISBN (Online):	979-8-3503-2244-6
ISSN:	1943-2895

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2023 IEEE 39th International Conference on Data Engineering Workshops (ICDEW) **ICDEW 2023**

Table of Contents

BDAFC 2023 - International Workshop on Big Data Analytics in Finance and Commerce

High-Dimensional Data from Financial Statements for a Bankruptcy Prediction Model	1
<i>Gianluca Gabrielli (University of Parma, Italy), Andrea Melioli (University of Parma, Italy), and Flavio Bertini (University of Parma, Italy)</i>	
Approximate Indexing for Top-k Queries over Massive Spatio-Textual Data Streams	8
<i>Hangjia Cen (University of Science and Technology of China, China), Xike Xie (University of Science and Technology of China, China), Xin Cao (University of New South Wales, Australia), and Jiali Weng (University of Science and Technology of China, China)</i>	
InvestLens: A Visual Analytics Approach to Inspecting the Dynamics of Venture Capital Investment Network	12
<i>Yun Tian (ShanghaiTech University, China), Liu Jia (ShanghaiTech University, China), Xinyi Zhang (ShanghaiTech University, China), Xingsheng Yang (ShanghaiTech University, China), Zhaoru Ke (ShanghaiTech University, China), Chenyang Zhang (University of Illinois at Urbana-Champaign, USA), Ling Liao (ABC-CA Fund Management Co., Ltd, China), Suting Hong (ShanghaiTech University, China), Haipeng Zhang (ShanghaiTech University, China), and Quan Li (ShanghaiTech University, China)</i>	
Prediction with Time-Series Mixer for the S&P500 Index	20
<i>Junyi Ye (New Jersey Institute of Technology), Jingyi Gu (New Jersey Institute of Technology), Ankan Dash (New Jersey Institute of Technology), Fadi P. Deek (New Jersey Institute of Technology), and Guiling Wang (New Jersey Institute of Technology)</i>	
Real-time Personalized Guest Experiences using Product Attribute Information on Vanilla Hardware	28
<i>Amit Pande (Target Corporation), Rahul Biswas (Target Corporation), Pushkar Chennu (Target Corporation), Harshith N (Target Corporation), and Anupama Joshi (Target Corporation)</i>	

HODL: The Hold of Reddit over the Stock Market	36
<i>A M Muntasir Rahman (New Jersey Institute of Technology, USA), Ajim Uddin (New Jersey Institute of Technology, USA), and Guiling Wang (New Jersey Institute of Technology, USA)</i>	

DBML 2023 - International Workshop on Databases and Machine Learning

Provenance-Based Explanations for Machine Learning (ML) Models	40
<i>Justin Turnau (University of Cincinnati, USA), Nkechi Akwari (University of Cincinnati, USA), Seokki Lee (University of Cincinnati, USA), and Dwarkesh Rajput (University of Cincinnati, USA)</i>	
Privacy-Preserving Data Federation for Trainable, Queryable and Actionable Data	44
<i>Stavroula Iatropoulou (Ubitech Limited, Cyprus), Theodora Anastasiou (Ubitech Limited, Cyprus), Sophia Karagiorgou (Ubitech Limited, Cyprus), Petros Petrou (Ubitech Limited, Cyprus), Dimitrios Alexandrou (Ubitech Limited, Cyprus), and Thanassis Bouras (Ubitech Limited, Cyprus)</i>	
Imputation of Missing Values in Training Data using Variational Autoencoder	49
<i>Xuerui Hong (Beijing Jiaotong University, China) and Shuang Hao (Beijing Jiaotong University, China)</i>	
COAX: Correlation-Aware Indexing	55
<i>Ali Hadian (Imperial College London), Behzad Ghaffari (Imperial College London), Taiyi Wang (Cambridge University), and Thomas Heinis (Imperial College London)</i>	
Efficient Index Learning via Model Reuse and Fine-Tuning	60
<i>Guanli Liu (The University of Melbourne, Australia), Jianzhong Qi (The University of Melbourne, Australia), Lars Kulik (The University of Melbourne, Australia), Kazuya Soga (The University of Melbourne, Australia), Renata Borovica-Gajic (The University of Melbourne, Australia), and Benjamin I. P. Rubinstein (The University of Melbourne, Australia)</i>	
A Fast Hybrid Spatial Index with External Memory Support	67
<i>Xinyu Su (The University of Melbourne, Australia), Jianzhong Qi (The University of Melbourne, Australia), and Egemen Tanin (The University of Melbourne, Australia)</i>	
Optimizing Machine Learning Inference Queries for Multiple Objectives	74
<i>Ziyu Li (Delft University of Technology, Netherlands), Mariette Schönfeld (Delft University of Technology, Netherlands), Rihan Hai (Delft University of Technology, Netherlands), Alessandro Bozzon (Delft University of Technology, Netherlands), and Asterios Katsifodimos (Delft University of Technology, Netherlands)</i>	
Unsupervised Intra-Domain Adaptation for Recommendation via Uncertainty Minimization	79
<i>Chenghao Chen (Alibaba Group, China), Jie Xiao (Alibaba Group, China), Jin Liu (Alibaba Group, China), Jie Zhang (Alibaba Group, China), Jia Jia (Alibaba Group, China), and Ning Hu (Alibaba Group, China)</i>	

SMDB - 14th International Workshop on Self-Managing Database Systems

SMDB 2023 Message and Committees	87
<i>Herodotos Herodotou (Cyprus University of Technology) and Yingjun Wu (Singularity Data Inc.)</i>	
SMDB 2023 Keynotes	91
<i>Anastasia Ailamaki (EPFL and Google, Inc.), Hanuma Kodavalla (Microsoft), Stratos Idreos (Harvard University), and Pedro Pedreira (Meta)</i>	
Towards Evaluating Stream Processing Autoscalers	95
<i>George Siachamis (Delft University of Technology), Job Kanis (Delft University of Technology), Wybe Koper (Delft University of Technology), Kyriakos Psarakis (Delft University of Technology), Marios Fragkoulis (Delivery Hero SE), Arie van Deursen (Delft University of Technology), and Asterios Katsifodimos (Delft University of Technology)</i>	
Towards a Signature Based Compression Technique for Big Data Storage	100
<i>Constantinos Costa (Rinnoco Ltd), Panos K. Chrysanthis (Rinnoco Ltd), Marios Costa (Rinnoco Ltd), Efstathios Stavrakis (Algolysis Ltd), and Nicolas Nicolaou (Algolysis Ltd)</i>	
Quantum Annealing Method for Dynamic Virtual Machine and Task Allocation in Cloud Infrastructures from Sustainability Perspective	105
<i>Valter Uotila (University of Helsinki, Finland) and Jiaheng Lu (University of Helsinki, Finland)</i>	
A Self-managed Marketplace for Sharing IoT Sensors	111
<i>Anas Dawod (Swinburne University of Technology), Dimitrios Georgakopoulos (Swinburne University of Technology), Prem Prakash Jayaraman (Swinburne University of Technology), and Panos K. Chrysanthis (University of Pittsburgh)</i>	

ASTRIDE 2023 - 1st International Workshop on Private, Secure, and Trustworthy IoT Data Management

Differential Privacy for Protecting Private Patterns in Data Streams	118
<i>He Gu (University of Oslo, Norway), Thomas Plagemann (University of Oslo, Norway), Maik Benndorf (University of Oslo, Norway), Vera Goebel (University of Oslo, Norway), and Boris Koldehofe (University of Groningen, The Netherlands)</i>	
Locally Private Streaming Data Release with Shuffling and Subsampling	125
<i>Xiaoyu Li (Kyoto University, Japan), Yang Cao (Hokkaido University, Japan), and Masatoshi Yoshikawa (Kyoto University, Japan)</i>	
Preserving Privacy in Image Database Through Bit-Planes Obfuscation	132
<i>Vishesh K. Tanwar (Missouri University of Science and Technology, USA), Ashish Gupta (Missouri University of Science and Technology, USA), Sanjay Madria (Missouri University of Science and Technology, USA), and Sajal K. Das (Missouri University of Science and Technology, USA)</i>	

Toward Compliance Implications and Security Objectives: A Qualitative Study	138
<i>Dmitry Prokhorenkov (Technical University of Munich, Germany)</i>	

DASC 2023 - Data-driven Smart Cities

Query-Driven Edge Node Selection in Distributed Learning Environments	146
<i>Tahani Aladwani (University of Glasgow, UK), Christos Anagnostopoulos (University of Glasgow, UK), Kostas Kolomvatsos (University of Thessaly), Ibrahim Alghamdi (Al-Baha University), and Fani Deligianni (University of Glasgow, UK)</i>	
Points-of-Interest from Mapillary Street-Level Imagery: A Dataset For Neighborhood Analytics	154
<i>Negin Zarbakhsh (University College Dublin, Ireland) and Gavin McArdle (University College Dublin, Ireland)</i>	
Dynamic Spatio-Temporal Access Queries using Semi-Supervised Regression	162
<i>Chris Conlan (University of Warwick, UK), Teddy Cunningham (University of Warwick, UK), and Hakan Ferhatosmanoglu (University of Warwick, UK)</i>	
Modeling Semantic Business Trajectories of Territories for Multidisciplinary Studies Through Controlled Vocabularies	170
<i>Muhammad Arslan (Université de Bourgogne, France) and Christophe Cruz (Université de Bourgogne, France)</i>	
A Smart Winter Road Maintenance Platform to Balance Monitoring Accuracy and Cost	178
<i>Janis Grabis (Riga Technical University, Latvia) and Lauma Jokste (Riga Technical University, Latvia)</i>	

Joint International Workshop on Big Data Management on Emerging Hardware and Data Management on Virtualized Active Systems (HardBD & Active'23)

An Empirical Performance Comparison Between Matrix Multiplication Join and Hash Join on GPUs	184
<i>Wenbo Sun (Delft University of Technology, The Netherlands), Asterios Katsifodimos (Delft University of Technology, The Netherlands), and Rihan Hai (Delft University of Technology, The Netherlands)</i>	
Adaptive Query Compilation with Processing-in-Memory	191
<i>Alexander Baumstark (Technische Universität Ilmenau, Germany), Muhammad Attahir Jibril (Technische Universität Ilmenau, Germany), and Kai-Uwe Sattler (Technische Universität Ilmenau, Germany)</i>	
Sailfish: Exploring Heterogeneous Query Acceleration on Discrete CPU-FPGA Architecture	198
<i>Xing Wei (State Key Laboratory of Mobile Network and Mobile Multimedia Technology, China; ZTE Corporation, China), Yaofeng Tu (State Key Laboratory of Mobile Network and Mobile Multimedia Technology, China; ZTE Corporation, China), Yinjun Han (State Key Laboratory of Mobile Network and Mobile Multimedia Technology, China; ZTE Corporation, China), Zhenghua Chen (ZTE Corporation, China), Xuecheng Qi (ZTE Corporation, China), and Daojun Hua (ZTE Corporation, China)</i>	

Personalized Semantic Matching for Web Search	205
<i>Kunlei Jing (Xi'an Jiaotong University, China), Fei Hao (The Hong Kong Polytechnic University, China), Xizi Zhang (Capital University of Economics and Trade, China), and Yu Zhou (Zhongnan University of Economics and Law, China)</i>	

6th International Workshop on Data Engineering Meets Intelligent Food and Cooking Recipes 2023 (DECOR 2023)

From the kitchen into the wild: Chances and challenges for data-driven AI in 21st century gastronomy	211
<i>Tarek R. Besold (Sony AI)</i>	
The Future of Eating Together: Exploring the Promises and Challenges of Digital Commensality in the Metaverse	212
<i>Alexandra Covaci (University of Kent), George Ghinea (Brunel University), and Frederic Andres (NII)</i>	
Data Engineering Challenges in Intelligent Food and Cooking Recipes	214
<i>Frederic Andres (National Institute of Informatics)</i>	
Feasting on Technology: The art of computer-enhanced serving	218
<i>Forouzan Golshani (College of Engineering, California State University, Long Beach, USA)</i>	
Author Index	221