



Proceedings of the VLDB Endowment

Volume 17, No. 11 – July 2024

Editors in Chief:

Meihui Zhang and Cyrus Shahabi

Associate Editors:

Alkis Polyzotis, Amol Deshpande, Angela Bonifati, Ashraf Abounaga, Ashwin Machanavajhala,
Beng Chin Ooi, Boris Glavic, Ce Zhang, Divy Agrawal, Eric Lo, Fatma Ozcan, Guoliang Li,
Jeffrey Xu Yu, Jian Pei, Jianliang Xu, Johannes Gehrke, K. Selçuk Candan, Kyuseok Shim, Li Xiong,
Magdalena Balazinska, Matthias Boehm, Melanie Herschel, Michael Böhlen,
Nikos Mamoulis, Pinar Tozun, Rachel Pottinger, Sharad Mehrotra, Surajit Chaudhuri, Tamer Özsu,
Tien Tuan Anh Dinh, Walid Aref, Wei Wang, Xiaokui Xiao, Yanyan Shen, Yongxin Tong, Zi Huang

Publication Editors:

Ju Fan, Yang Cao, Xiaou Ding

PVLDB – Proceedings of the VLDB Endowment

Volume 17, No. 11, July 2024.

All papers published in this issue will be presented at the 50th International Conference on Very Large Data Bases, Guangzhou, China, 2024.

Copyright 2024 VLDB Endowment

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>. For any use beyond those covered by this license, obtain permission by emailing info@vldb.org.

Volume 17, Number 11, July 2024

Pages i – x and 2681 - 3719

ISSN 2150-8097

Available at: <http://www.pvldb.org> and <https://dl.acm.org/journal/pvldb>

TABLE OF CONTENTS

Front Matter

Copyright Notice	i
Table of Contents	ii
PVLDB Organization and Review Board – Vol. 17	iv

Research Papers

Robust Best Point Selection under Unreliable User Feedback.....	2681
<i>Qixu Chen, Raymond Chi-Wing Wong</i>	
Towards Optimal Transaction Scheduling.....	2694
<i>Audrey Cheng, Aaron Kabcenell, Jason Chan, Xiao Shi, Peter Bailis, Natacha Crooks, Ion Stoica</i>	
QCore: Data-Efficient, On-Device Continual Calibration for Quantized Models.....	2708
<i>David Campos, Bin Yang, Tung Kieu, Miao Zhang, Chenjuan Guo, Christian S. Jensen</i>	
Efficient Algorithms for Pseudoarboricity Computation in Large Static and Dynamic Graphs.....	2722
<i>Yalong Zhang, Ronghua Li, Qi Zhang, Hongchao Qin, Lu Qin, Guoren Wang</i>	
RoarGraph: A Projected Bipartite Graph for Efficient Cross-Modal Approximate Nearest Neighbor Search	2735
<i>Meng Chen, Kai Zhang, Zhenying He, Yinan Jing, X. Sean Wang</i>	
Combining Small Language Models and Large Language Models for Zero-Shot NL2SQL.....	2750
<i>Ju Fan, Zihui Gu, Songyue Zhang, Yuxin Zhang, Zui Chen, Lei Cao, Guoliang Li, Samuel Madden, Xiaoyong Du, Nan Tang</i>	
D3-GNN: Dynamic Distributed Dataflow for Streaming Graph Neural Networks.....	2764
<i>Rustam Guliyev, Aparajita Haldar, Hakan Ferhatosmanoglu</i>	
Distance-based Outlier Query Optimization in Apache IoTDB	2778
<i>Yunxiang Su, Shaoxu Song, Xiangdong Huang, Chen Wang, Jianmin Wang</i>	
TC-Match: Fast Time-constrained Continuous Subgraph Matching	2791
<i>Jianye Yang, Sheng Fang, Zhaoquan Gu, Ziyi Ma, Xuemin Lin, Zhihong Tian</i>	
Automating the Enterprise with Foundation Models.....	2805
<i>Michael Wornow, Avanika Narayan, Krista Opsahl-Ong, Quinn Mcintyre, Nigam Shah, Christopher Re</i>	
Efficient Index for Temporal Core Queries over Bipartite Graphs	2813
<i>Anxin Tian, Alexander Zhou, Yue Wang, Xun Jian, Lei Chen</i>	
Uldp-FL: Federated Learning with Across Silo User-Level Differential Privacy	2826
<i>Fumiyuki Kato, Li Xiong, Shun Takagi, Yang Cao, Masatoshi Yoshikawa</i>	
Evolution Forest Index: Towards Optimal Temporal k -Core Component Search via Time-Topology Isomorphic Computation	2840
<i>Junyong Yang, Ming Zhong, Yuanyuan Zhu, Tieyun Qian, Mengchi Liu, Jeffrey Xu Yu</i>	

Eliminating Data Processing Bottlenecks in GNN Training over Large Graphs via Two-level Feature Compression	2854
<i>Yuxin Ma, Ping Gong, Tianming Wu, Jiawei Yi, Chengru Yang, Cheng Li, Qirong Peng, Guiming Xie, Yongcheng Bao, Haifeng Liu, Yinlong Xu</i>	
Towards Systematic Index Dynamization	2867
<i>Douglas B. Rumbaugh, Dong Xie, Zhuoyue Zhao</i>	
Ensemble Clustering based on Meta-Learning and Hyperparameter Optimization	2880
<i>Dennis Treder-Tschechlov, Manuel Fritz, Holger Schwarz, Bernhard Mitschang</i>	
Efficient Stochastic Routing in Path-Centric Uncertain Road Networks.....	2893
<i>Chenjuan Guo, Ronghui Xu, Bin Yang, Yuan Ye, Tung Kieu, Yan Zhao, Christian S. Jensen</i>	
Transforming Property Graphs.....	2906
<i>Angela Bonifati, Filip Murlak, Yann Ramusat</i>	
Are Large Language Models a Good Replacement of Taxonomies?	2919
<i>Yushi Sun, Xin Hao, Kai Sun, Yifan Ethan Xu, Xiao Yang, Xin Luna Dong, Nan Tang, Lei Chen</i>	
Efficient Algorithms for Density Decomposition on Large Static and Dynamic Graphs	2933
<i>Yalong Zhang, Ronghua Li, Qi Zhang, Hongchao Qin, Guoren Wang</i>	
Efficient Maximal Motif-Clique Enumeration over Large Heterogeneous Information Networks	2946
<i>Yingli Zhou, Yixiang Fang, Chenhao Ma, Tianci Hou, Xin Huang</i>	
OUTRE: An OUT-of-core De-REdundancy GNN Training Framework for Massive Graphs within A Single Machine.....	2960
<i>Zeang Sheng, Wentao Zhang, Yangyu Tao, Bin Cui</i>	
On Reducing Space Amplification with Multi-Column Compaction in Apache IoTDB.....	2974
<i>Chenguang Fang, Zijie Chen, Shaoxu Song, Xiangdong Huang, Chen Wang, Jianmin Wang</i>	
AutoTSAD: Unsupervised Holistic Anomaly Detection for Time Series Data	2987
<i>Sebastian Schmidl, Felix Naumann, Thorsten Papenbrock</i>	
Enabling Window-Based Monotonic Graph Analytics with Reusable Transitional Results for Pattern-Consistent Queries	3003
<i>Zheng Chen, Feng Zhang, Yang Chen, Xiaokun Fang, Guanyu Feng, Xiaowei Zhu, Wenguang Chen, Xiaoyong Du</i>	
When Amnesia Strikes: Understanding and Reproducing Data Loss Bugs with Fault Injection	3017
<i>Maria Ramos, João Azevedo, Kyle Kingsbury, José Pereira, Tânia Esteves, Ricardo Macedo, João Paulo</i>	
PriPL-Tree: Accurate Range Query for Arbitrary Distribution under Local Differential Privacy.....	3031
<i>Leixia Wang, Qingqing Ye, Haibo Hu, Xiaofeng Meng</i>	
Win-Win: On Simultaneous Clustering and Imputing over Incomplete Data.....	3045
<i>Yu Sun, Jingyu Zhu, Xiao Xu, Xian Xu, Yuyao Sun, Shaoxu Song, Xiang Li, Xiaojie Yuan</i>	
HRNet: Differentially Private Hierarchical and Multi-Resolution Network for Human Mobility Data Synthesization.....	3058
<i>Shun Takagi, Li Xiong, Fumiyuki Kato, Yang Cao, Masatoshi Yoshikawa</i>	

Efficiently Mitigating the Impact of Data Drift on Machine Learning Pipelines	3072
<i>Sijie Dong, Qitong Wang, Soror Sahri, Themis Palpanas, Divesh Srivastava</i>	
PCSP: Efficiently Answering Label-Constrained Shortest Path Queries in Road Networks.....	3082
<i>Libin Wang, Raymond Chi-Wing Wong</i>	
Efficient Enumeration of Recursive Plans in Transformation-based Query Optimizers.....	3095
<i>Amela Fejza, Pierre Genevès, Nabil Layaida</i>	
Enriching Relations with Additional Attributes for ER	3109
<i>Mengyi Yan, Wenfei Fan, Yaoshu Wang, Min Xie</i>	
Enhancing Accuracy for Super Spreader Identification in High-Speed Data Streams.....	3124
<i>Haibo Wang</i>	
Privately Answering Queries on Skewed Data via Per-Record Differential Privacy	3138
<i>Jeremy Seeman, William Sexton, David Pujol, Ashwin Machanavajhala</i>	
Agile-Ant: Self-managing Distributed Cache Management for Cost Optimization of Big Data Applications	3151
<i>Hani Al-Sayeh, Muhammad Attahir Jibril, Kai-Uwe Sattler</i>	
Complex Event Recognition with Symbolic Register Transducers	3165
<i>Elias Alevizos, Alexander Artikis, Georgios Paliouras</i>	
HAICChart: Human and AI Paired Visualization System	3178
<i>Yupeng Xie, Yuyu Luo, Guoliang Li, Nan Tang</i>	
A Sampling-based Framework for Hypothesis Testing on Large Attributed Graphs.....	3192
<i>Yun Wang, Chrysanthi Kosyfaki, Sihem Amer-Yahia, Reynold Cheng</i>	
LLM-PBE: Assessing Data Privacy in Large Language Models	3201
<i>Qinbin Li, Junyuan Hong, Chulin Xie, Jeffrey Tan, Rachel Xin, Junyi Hou, Xavier Yin, Zhun Wang, Dan Hendrycks, Zhangyang Wang, Bo Li, Bingsheng He, Dawn Song</i>	
Robust Join Processing with Diamond Hardened Joins.....	3215
<i>Altan Birler, Alfons Kemper, Thomas Neumann</i>	
DARKER: Efficient Transformer with Data-driven Attention Mechanism for Time Series	3229
<i>Rundong Zuo, Guozhong Li, Rui Cao, Byron Choi, Jianliang Xu, Sourav S Bhowmick</i>	
Efficient Maximal Frequent Group Enumeration in Temporal Bipartite Graphs.....	3243
<i>Yanping Wu, Renjie Sun, Xiaoyang Wang, Dong Wen, Ying Zhang, Lu Qin, Xuemin Lin</i>	
Optimizing Video Queries with Declarative Clues	3256
<i>Daren Chao, Yueting Chen, Nick Koudas, Xiaohui Yu</i>	
Fainder: A Fast and Accurate Index for Distribution-Aware Dataset Search	3269
<i>Lennart Behme, Sainyam Galhotra, Kaustubh Beedkar, Volker Markl</i>	
nsDB: Architecting the Next Generation Database by Integrating Neural and Symbolic Systems.....	3283
<i>Ye Yuan, Bo Tang, Tianfei Zhou, Zhiwei Zhang, Jianbin Qin</i>	
Two Birds With One Stone: Designing a Hybrid Cloud Storage Engine for HTAP.....	3290
<i>Tobias Schmidt, Dominik Durner, Viktor Leis, Thomas Neumann</i>	

DDS: DPU-optimized Disaggregated Storage	3304
<i>Qizhen Zhang, Philip A. Bernstein, Badrish Chandramouli, Jiasheng Hu, Yiming Zheng</i>	
The Dawn of Natural Language to SQL: Are We Fully Ready?	3318
<i>Boyan Li, Yuyu Luo, Chengliang Chai, Guoliang Li, Nan Tang</i>	
Counterfactual Explanation of the Shapley Value in Data Coalitions	3332
<i>Michelle Si, Jian Pei</i>	
Searching Data Lakes for Nested and Joined Data	3346
<i>Yi Zhang, Peter Baile Chen, Zackary G. Ives</i>	
Efficient Betweenness Centrality Computation over Large Heterogeneous Information Networks	3360
<i>Xinrui Wang, Wang Yiran, Xuemin Lin, Jeffrey Xu Yu, Hong Gao, Xiuzhen Cheng, Dongxiao Yu</i>	
The Holon Approach for Simultaneously Tuning Multiple Components in a Self-Driving Database Management System with Machine Learning via Synthesized Proto-Actions	3373
<i>William Zhang, Wan Shen Lim, Matthew Butrovich, Andrew Pavlo</i>	
DynaHB: A Communication-Avoiding Asynchronous Distributed Framework with Hybrid Batches for Dynamic GNN Training	3388
<i>Zhen Song, Yu Gu, Qing Sun, Tianyi Li, Yanfeng Zhang, Yushuai Li, Christian S. Jensen, Ge Yu</i>	
Efficient kNN Search in Public Transportation Networks	3402
<i>Qingshuai Feng, Junhua Zhang, Wenjie Zhang, Lu Qin, Ying Zhang, Xuemin Lin</i>	
LITS: An Optimized Learned Index for Strings	3415
<i>Yifan Yang, Shimin Chen</i>	
OLAP on Modern Chiplet-Based Processors	3428
<i>Alessandro Fogli, Bo Zhao, Peter Pietzuch, Maximilian Bandle, Jana Giceva</i>	
Bf-Tree: A Modern Read-Write-Optimized Concurrent Larger-Than-Memory Range Index	3442
<i>Xiangpeng Hao, Badrish Chandramouli</i>	
Partition, Don't Sort! Compression Boosters for Cloud Data Ingestion Pipelines	3456
<i>Patrick Hansert, Sebastian Michel</i>	
Chameleon: Foundation Models for Fairness-aware Multi-modal Data Augmentation to Enhance Coverage of Minorities	3470
<i>Mahdi Erfanian, H. V. Jagadish, Abolfazl Asudeh</i>	
DAFDiscover: Robust Mining Algorithm for Dynamic Approximate Functional Dependencies on Dirty Data	3484
<i>Xiaoou Ding, Yixing Lu, Hongzhi Wang, Chen Wang, Yida Liu, Jianmin Wang</i>	
Hardware-Efficient Data Imputation through DBMS Extensibility	3497
<i>Hubert Mohr-Daurat, Georgios Theodorakis, Holger Pirk</i>	
Generating Succinct Descriptions of Database Schemata for Cost-Efficient Prompting of Large Language Models	3511
<i>Immanuel Trummer</i>	
Saving Money for Analytical Workloads in the Cloud	3524
<i>Tapán Srivastava, Raul Castro Fernandez</i>	

ReCG: Bottom-Up JSON Schema Discovery Using a Repetitive Cluster-and-Generalize Framework	3538
<i>Joohyung Yun, Byungchul Tak, Wook-Shin Han</i>	
Optimizing Collections of Bloom Filters within a Space Budget.....	3551
<i>Gabriel Mersy, Zhuo Wang, Stavros Sintos, Sanjay Krishnan</i>	
A Spark Optimizer for Adaptive, Fine-Grained Parameter Tuning	3565
<i>Chenghao Lyu, Qi Fan, Philippe Guyard, Yanlei Diao</i>	
Texera: A System for Collaborative and Interactive Data Analytics Using Workflows	3580
<i>Zuozhi Wang, Yicong Huang, Shengquan Ni, Avinash Kumar, Sadeem Alsudais, Xiaozhen Liu, Xinyuan Lin, Yunyan Ding, Chen Li</i>	
Efficient Validation of SHACL Shapes with Reasoning.....	3589
<i>Jin Ke, Zenon Zacouris, Maribel Acosta</i>	
QED: A Powerful Query Equivalence Decider for SQL	3602
<i>Shuxian Wang, Sicheng Pan, Alvin Cheung</i>	
Index Advisors on Quantum Platforms.....	3615
<i>Manish Kesarwani, Jayant R. Haritsa</i>	
Blueprinting the Cloud: Unifying and Automatically Optimizing Cloud Data Infrastructures with BRAD	3629
<i>Geoffrey X. Yu, Ziniu Wu, Ferdi Kossmann, Tianyu Li, Markos Markakis, Amadou Ngom, Samuel Madden, Tim Kraska</i>	
Aleph Filter: To Infinity in Constant Time.....	3644
<i>Niv Dayan, Ioana-Oriana Bercea, Rasmus Pagh</i>	
RUSH: Real-time Burst Subgraph Discovery in Dynamic Graphs.....	3657
<i>Yuhang Chen, Jiaxin Jiang, Shixuan Sun, Bingsheng He, Min Chen</i>	
A Benchmark Study of Deep-RL Methods for Maximum Coverage Problems over Graphs	3666
<i>Zhicheng Liang, Yu Yang, Xiangyu Ke, Xiaokui Xiao, Yunjun Gao</i>	
Hit the Gym: Accelerating Query Execution to Efficiently Bootstrap Behavior Models for Self-Driving Database Management Systems.....	3680
<i>Wan Shen Lim, Lin Ma, William Zhang, Matthew Butrovich, Samuel Arch, Andrew Pavlo</i>	
Why TPC Is Not Enough: An Analysis of the Amazon Redshift Fleet.....	3694
<i>Alexander van Renen, Dominik Horn, Pascal Pfeil, Kapil Vaidya, Wenjian Dong, Murali Narayanaswamy, Zhengchun Liu, Gaurav Saxena, Andreas Kipf, Tim Kraska</i>	
Efficient k -Clique Count Estimation with Accuracy Guarantee	3707
<i>Lijun Chang, Rashmika Gamage, Jeffrey Xu Yu</i>	

PVLDB ORGANIZATION AND REVIEW BOARD - Vol. 17

Editors in Chief of PVLDB

Meihui Zhang (Beijing Institute of Technology)
Cyrus Shahabi (University of Southern California)

Associate Editors of PVLDB

Alkis Polyzotis (Databricks)
Amol Deshpande (University of Maryland at College Park)
Angela Bonifati (Lyon 1 University)
Ashraf Aboulnaga (Qatar Computing Research Institute, HBKU)
Ashwin Machanavajjhala (Duke)
Beng Chin Ooi (NUS)
Boris Glavic (Illinois Institute of Technology)
Ce Zhang (ETH)
Divy Agrawal (University of California, Santa Barbara)
Eric Lo (Chinese University of Hong Kong)
Fatma Ozcan (Google)
Guoliang Li (Tsinghua University)
Jeffrey Xu Yu (Chinese University of Hong Kong)
Jian Pei (Simon Fraser University)
Jianliang Xu (Hong Kong Baptist University)
Johannes Gehrke (Microsoft)
K. Selçuk Candan (Arizona State University)
Kyuseok Shim (Seoul National University)
Li Xiong (Emory University)
Magdalena Balazinska (UW)
Matthias Boehm (Technische Universität Berlin)
Melanie Herschel (Universität Stuttgart)
Michael Böhlen (University of Zurich)
Nikos Mamoulis (University of Ioannina)
Pinar Tozun (IT University of Copenhagen)
Rachel Pottinger (Univ. of British Columbia)
Sharad Mehrotra (U.C. Irvine)
Surajit Chaudhuri (Microsoft)

Tamer Özsu (University of Waterloo)
Tien Tuan Anh Dinh (Deakin University)
Walid Aref (Purdue University)
Wei Wang (ByteDance)
Xiaokui Xiao (National University of Singapore)
Yanyan Shen (Shanghai Jiao Tong University)
Yongxin Tong (Beihang University)
Zi Huang (University of Queensland)

Publication Editors

Ju Fan (Renmin University of China)
Yang Cao (Tokyo Institute of Technology)
Xiaou Ding (Harbin Institute of Technology)

PVLDB Managing Editor

Wolfgang Lehner (Dresden University of Technology)

PVLDB Advisory Board

Vanessa Braganholo (Universidade Federal Fluminense)
Sourav S Bhowmick (Nanyang Technological University)
Torsten Grust (University of Tuebingen)
Xin Luna Dong (Facebook)
Fatma Ozcan (Google)
Lei Chen (Hong Kong University of S&T)
Juliana Freire (New York University)
Graham Cormode (University of Warwick)
Divesh Srivastava (AT&T Labs-Research)
Felix Naumann (HPI)
Georgia Koutrika (Athena Research Center)
Jun Yang (Duke University)
Meihui Zhang (Beijing Institute of Technology)
Cyrus Shahabi (University of Southern California)
Nesime Tatbul (Intel Labs and MIT)
Themis Palpanas (Universite Paris Cite)

Review Board

Abolfazl Asudeh (University of Illinois Chicago)
Aditya Parameswaran (University of California, Berkeley)
Ahmed S. Abdelhamid (Purdue University)
Ahmed Eldawy (University of California, Riverside)
Ahmed El-Roby (Carleton University)
Ahmed Mahmood (Google)
Alberto Lerner (University of Fribourg, Switzerland)
Alexander Thomson (Google)
Amr Magdy (University of California Riverside)
Andreas Züfle (Emory University)
Angelos Christos Anadiotis (Oracle)
Anja Gruenheid (Microsoft)
Anthony Tung (National U. of Singapore)
Anton Dignös (Free University of Bozen-Bolzano, Italy)
Arijit Khan (Aalborg University)
Avrilia Floratou (Microsoft)
Baihua Zheng (Singapore Management University)
Bailu Ding (Microsoft Research)
Berthold Reinwald (IBM Research-Almaden)
Bin Yang (East China Normal University)
Bingsheng He (National University of Singapore)
Bolin Ding (Data Analytics and Intelligence Lab, Alibaba Group)
Brandon Haynes (Microsoft Gray Systems Lab)
Chao Zhang (University of Waterloo)
Cheng Long (Nanyang Technological University)
Chengfei Liu (Swinburne University of Technology)
Chengkai Li (The University of Texas at Arlington)
Chengliang Chai (Beijing Institute of Technology)
Chrysanthi Kosyfaki (University of Ioannina)
Chunwei Liu (MIT)
Cong Yan (Microsoft research)
Daisy Zhe Wang (University of Florida)
Dan Kifer (Pennsylvania State Univ., USA)
Dan Lin (Vanderbilt University)
Daniel Kang (UIUC)
Demetrios Zeinalipour-Yazti (University of Cyprus)
Dimitris Papadias (HKUST)
Dong Deng (Rutgers University - New Brunswick)
Dong Wen (University of New South Wales)
Dong Xie (Penn State University)
Dongxiang Zhang (Zhejiang University)
Dumitrel Loghin (National University of Singapore)
Egemen Tanin (University of Melbourne)
El Kindi Rezig (Massachusetts Institute of Technology)
Elena Ferrari (University of Insubria, Varese)
Eser Kandogan (Megagon Labs)
Essam Mansour (Concordia University)
Fan Zhang (Guangzhou University)
Fatemeh Nargesian (University of Rochester)
Fei Chiang (McMaster University)
Feng Zhang (Renmin University of China)
Florin Rusu (UC Merced)
Gabriel Ghinita (Hamad Bin Khalifa University)
Gao Cong (Nanyang Technological University)
George Fakas (Uppsala University)
Haibo Hu (Hong Kong Polytechnic University)
Holger Pirk (Imperial College)
Hong Cheng (Chinese University of Hong Kong)

Hongzhi Wang (Harbin Institute of Technology)
Hua Lu (Roskilde University)
Huanchen Zhang (Tsinghua University)
Huiping Cao (New Mexico State University)
Ibrahim Sabek (MIT)
Ilaria Bartolini (University of Bologna)
Jana Giceva (TU Munich)
Jennie Rogers (Northwestern University)
Jia Zou (Arizona State University)
Jian Lou (Zhejiang University)
Jiangshan Yu (Monash University)
Jianguo Wang (Purdue University)
Jiannan Wang (Simon Fraser University)
Jianqiu Xu (Nanjing University of Aeronautics and Astronautics)
Jianxin Li (Deakin University)
Jieming Shi (The Hong Kong Polytechnic University)
Jin Wang (Megagon Labs)
Jinfei Liu (Zhejiang University)
Johes Bater (Tufts University)
John Liagouris (Boston University)
Jonathan Goldstein (Microsoft)
Ju Fan (Renmin University of China)
Juhua Hu (University of Washington)
Kai Wang (Shanghai Jiao Tong University)
Kangfei Zhao (Beijing Institute of Technology)
Karima Echihabi (Mohammed VI Polytechnic University)
Katja Hose (TU Wien)
Khuzaima Daudjee (University of Waterloo)
Kyoungmin Kim (POSTECH)
Lawrence Benson (HPI, University of Potsdam)
Lei Chen (Hong Kong University of Science and Technology)
Lei Zou (Peking University)
Leong Hou U (University of Macau)
Lin Ma (University of Michigan)
Linyang Chu (McMaster University)
Liyue Fan (UNC Charlotte)
Lu Chen (Zhejiang University)
Luigi Bellomarini (Banca d'Italia)
Madelon Hulsebos (University of Amsterdam)
Manolis Terrovitis (IMIS, Athena RC)
Marco Patella (University of Bologna)
Mario Nascimento (Northeastern University)
Matteo Lissandrini (Aalborg University)
Matthias Renz (University of Kiel)
Michael Hay (Colgate University & Tumult Labs)
mingjie tang (Ant Financial)
Mirek Riedewald (Northeastern University)
Mohamed S. Hassan (Google)
Mohamed Mokbel (University of Minnesota - Twin Cities)
Mohammad Javad Amiri (University of Pennsylvania)
Mostafa Milani (The University of Western Ontario)
Mourad OUZZANI (Qatar Computing Research Institute, HBKU)
Nesime Tatbul (Intel Labs and MIT)
Norman May (SAP SE)
Oliver A Kennedy (University at Buffalo, SUNY)

Panagiotis Bouros (Johannes Gutenberg University Mainz)
 Papotti Paolo (EURECOM)
 Patrick Damme (Technische Universität Berlin)
 Peng Peng (Hunan University)
 Philippe Bonnet (IT Univ Copenhagen, Denmark)
 Pinar Karagoz (METU, Turkey)
 Prashant Pandey (University of Utah)
 Primal Pappachan (Penn State University)
 Qichen Wang (Hong Kong Baptist University)
 Qing Liu (Zhejiang University)
 Qun Chen (Northwestern Polytechnical University)
 Renata Borovica-Gajic (University of Melbourne)
 Rihan Hai (TU Delft)
 Ritesh Ahuja (Oracle Labs)
 Roger Zimmermann (NUS)
 Ronghua Li (Beijing Institute of Technology)
 Sai Wu (Zhejiang Univ)
 Sanjay Krishnan (UChicago)
 Senjuti Basu Roy (NJIT)
 Seokki Lee (University of Cincinnati)
 Shantanu Sharma (New Jersey Institute of Technology)
 Shaofeng Cai (National University of Singapore)
 Shaoxu Song (Tsinghua University)
 Shuai Ma (Beihang University)
 Shuang Hao (Beijing Jiaotong University)
 Sibowang (The Chinese University of Hong Kong)
 Stefania Dumbrava (ENSIIE)
 Stefano Paraboschi (Universita' degli Studi di Bergamo)
 Sujaya Maiyya (University of Waterloo)
 Tarique Siddiqui (Microsoft Research)
 Thanana Ghanem (Metro State University)
 Thang Dinh (VCU)
 Themis Palpanas (Universite Paris Cite)
 Thomas Neumann (TUM)
 Tianhao Wang (University of Virginia)
 Tianzheng Wang (Simon Fraser University)
 Tieying Zhang (Bytedance)
 Tristan Allard (Univ Rennes, CNRS, IRISA)
 Umar Farooq Minhas (Apple)
 Utku Sirin (Harvard University)
 Viktor Leis (Technische Universität München)
 Vincenzo Gulisano (Chalmers University of Technology)
 Vraj Shah (IBM Research)
 Wang-Chien Lee (Pennsylvania State University, USA)
 WEI LU (Renmin University of China)

Wei Wang (Hong Kong University of Science and Technology (Guangzhou))
 Wei-Shinn Ku (Auburn University)
 Wenchao Zhou (Alibaba Group)
 Wendy Hui Wang (Stevens Institute of Technology)
 Xiang Lian (Kent State University)
 Xiang Zhao (National University of Defence Technology)
 Xiangyao Yu (University of Wisconsin-Madison)
 Xiao Hu (Duke University)
 Xiao Hu (University of Waterloo)
 Xiaochun Yang (Northeastern University)
 Xiaofang Zhou (The Hong Kong University of Science and Technology)
 Xiaofei Zhang (University of Memphis)
 Xiaohui Yu (York University)
 Xiaoli Wang (Xiamen University)
 Xin Huang (Hong Kong Baptist University)
 Xin Wang (Tianjin University)
 Xingquan Zhu (Florida Atlantic University)
 Yanfeng Zhang (Northeastern University)
 Yang Cao (Tokyo Institute of Technology)
 Yannis Chronis (Google)
 Yao Lu (Microsoft Research)
 Ye Yuan (Beijing Institute of Technology)
 Yeye He (Microsoft Research)
 Ying Zhang (University of Technology Sydney)
 Yingxia Shao (BUPT)
 Yu Yang (City University of Hong Kong)
 Yuhao Zhang (University of California, San Diego)
 Yuncheng Wu (National University of Singapore)
 Yunjun Gao (Zhejiang University)
 Yuval Moskovitch (Ben Gurion University)
 Yuxiang Zeng (Beihang University)
 Zhaojing Luo (National University of Singapore)
 Zhengjie Miao (Duke University)
 Zhichao Cao (Arizona State University)
 Zhifeng Bao (RMIT University)
 Zhiwei Zhang (Beijing Institute of Technology)
 Zhongle Xie (Zhejiang University)
 Zhuoyue Zhao (University at Buffalo - SUNY)
 Ziawasch Abedjan (Leibniz Universität Hannover)
 Ziliang Lai (Chinese University of Hong Kong)
 Zimu Zhou (City University of Hong Kong)

LETTER FROM THE EDITORS IN CHIEF

It is our pleasure to present the eleventh issue of Volume 17 of PVLDB (Proceedings of the VLDB). PVLDB is dedicated to showcasing original research papers that encompass a wide spectrum of subjects within the area of data and information management. Our coverage spans from fundamental theoretical principles and cutting-edge system architectures to innovative models, techniques, novel applications, and the comprehensive assessment and deployment of large-scale solutions. In our research track, we feature four equally significant categories of papers: (a) regular research, (b) scalable data science (SDS), (c) experiment, analysis & benchmark (EA&B), and (d) vision papers.

The eleventh issue of PVLDB's Volume 17 includes 78 papers, spanning the topics of Graph and Network Data; Database Engines; Information Integration and Data Quality; Machine Learning, AI and Databases; Novel Database Architectures; Data Privacy and Security; Distributed Database Systems; Data Mining and Analytics; Specialized and Domain-Specific Data Management; Data models and query languages; Provenance and Workflows; User Interfaces; and Text and Semi-Structured Data. Four topics stood out, the most popular ones in this issue: Graph and Network Data (16 papers), Database Engines (11 papers), Information Integration and Data Quality (10 papers), and Machine Learning, AI, and Databases (9 papers).

Out of the 78 papers, five are in the experiment, analysis & benchmark category, four is in the scalable data science category, two in the vision category, and the rest are regular research papers. Twenty-one papers were accepted after a shepherding round, and the others were accepted after revision.

PVLDB is committed to providing valuable and constructive feedback through a rigorous review process. All submissions undergo meticulous peer review by a team of accomplished Associate Editors and dedicated reviewers. Each paper receives comprehensive evaluation from a minimum of three reviewers, along with the oversight of an Associate Editor. During a three-week discussion phase, reviewers engage in a thorough exchange of perspectives, ultimately converging on a consensus, which is summarized in a meta-review. Some submissions may proceed to a revision phase, affording authors a three-month window to refine their work for subsequent review cycles.

This issue is the result of all the work put in by the authors as well as the great commitment and effort of our associate editors and reviewers as well as our proceedings chairs.

Meihui Zhang and Cyrus Shahabi
Editors-in-Chief of PVLDB Vol. 17
Program Chairs for VLDB 2024