Proceedings of the VLDB Endowment

Volume 18, No. 7 – March 2025

Editors in Chief:

Themis Palpanas and Nesime Tatbul

Associate Editors:

Walid G. Aref, Manos Athanassoulis, Carsten Binnig, Spyros Blanas, Matthias Boehm, Angela Bonifati, K. Selcuk Candan, Lei Cao, Raul Castro Fernandez, Lei Chen, Shimin Chen, Yi Chen, Reynold Cheng, Alvin Cheung, Sudipto Das, Niv Dayan, Antonis Deligiannakis, Jens Dittrich, Xin Luna Dong, Karima Echihabi, Alan Fekete, Avrilia Floratou, Jana Giceva, Katja Hose, H. V. Jagadish, Panos Kalnis, Georgia Koutrika, Eric Lo, Nikos Mamoulis, Stefan Manegold, Ioana Manolescu, Norman May, Umar Farooq Minhas, Fatemeh Nargesian, Beng Chin Ooi, Fatma Ozcan, Tamer Ozsu, Tilmann Rabl, Mirek Riedewald, Jennie Rogers, Alkis Simitsis, Letizia Tanca, Nan Tang, Yuanyuan Tian, Yongxin Tong, Pinar Tozun, Yannis Velegrakis, Matthias Weidlich, Steven E. Whang, Raymond Chi-Wing Wong

Publication Editors:

Xiaoou Ding, Subhadeep Sarkar, Giovanni Simonini

PVLDB – Proceedings of the VLDB Endowment

Volume 18, No. 7, March 2025.

All papers published in this issue will be presented at the 51st International Conference on Very Large Data Bases, London, United Kingdom, 2025.

Copyright 2025 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 18, Number 7, March 2025

Pages i – viii and 2005-2307

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. 18 iv

Research Papers

Efficient Maintenance of 2-Hop Labeling Index on Dynamic Small-World Graphs 2005
Yuanyuan Zeng, Yixiang Fang, Kun Chen, Yangfan Li, Chenhao Ma

Vive la Différence: Practical Diff Testing of Stateful Applications 2018
Kexin Zhu, Michael Whittaker, Srdjan Petrovic, Robert Grandl, Sanjay Ghemawat

GREAT: Generalized Reservoir Sampling based Triangle Counting Estimation over Streaming Graphs . 2031
Siyue Wu, Dingming Wu, Sinhong Cheuk, Tsz Nam Chan, Kezhong Lu

Efficient Discovery of Relaxed Functional Dependencies 2044
Mengran Li, Zijing Tan, Honghui Yang, Shuai Ma

SimRN: Trajectory Similarity Learning in Road Networks based on Distributed Deep Reinforcement Learning 2057
Danlei Hu, Yilin Li, Lu Chen, Ziquan Fang, Yushuai Li, Yunjun Gao, Tianyi Li

BIRDIE: Natural Language-Driven Table Discovery Using Differentiable Search Index 2070
Yuxiang Guo, Zhonghao Hu, Yuren Mao, Baihua Zheng, Yunjun Gao, Mingwei Zhou

SkyStore: Cost-Optimized Object Storage Across Regions and Clouds 2084
Shu Liu, Xiangxi Mo, Moshik Hershcovitch, Henric Zhang, Audrey Cheng, Guy Girmonsky, Gil Vernik, Michael Factor, Tiemo Bang, Soujanya Ponnapalli, Natacha Crooks, Joseph Gonzalez, Danny Harnik, Ion Stoica

The Power of Constraints in Natural Language to SQL Translation 2097
Tonghui Ren, Chen Ke, Yuankai Fan, Yinan Jing, Zhenying He, Kai Zhang, X. Sean Wang

ACE: A Cardinality Estimator for Set-Valued Queries 2112
Yufan Sheng, Xin Cao, Kaiqi Zhao, Yixiang Fang, Jianzhong Qi, Wenjie Zhang, Christian S. Jensen

Accio: Bolt-on Query Federation 2126
Xiaoying Wang, Jiannan Wang, Tianzheng Wang, Yong Zhang

Falcon: Advancing Asynchronous BFT Consensus for Lower Latency and Enhanced Throughput 2136
Xiaohai Dai, Chaozheng Ding, Wei Li, Jiang Xiao, Bolin Zhang, Chen Yu, Albert Zomaya, Hai Jin

Scalable Pre-Training of Compact Urban Spatio-Temporal Predictive Models on Large-Scale Multi-Domain Data 2149
Jindong Han, Hao Wang, Hui Xiong, Hao Liu

HADES: Range-Filtered Private Aggregation on Public Data 2159
Xiaoyuan Liu, Ni Trieu, Trinabh Gupta, Ishtiyaque Ahmad, Dawn Song

Optimized Batch Prompting for Cost-effective LLMs 2172
Zhaoxuan Ji, Xinlu Wang, Zhaojing Luo, Zhongle Xie, Meihui Zhang

Truss Decomposition in Hypergraphs 2185
Hongchao Qin, Guang Zeng, Ronghua Li, Longlong Lin, Ye Yuan, Guoren Wang

Optimal Sharding for Scalable Blockchains with Deconstructed SMR 2198
Jianting Zhang, Zhongtang Luo, Raghavendra Ramesh, Aniket Kate

Auto-Prep: Holistic Prediction of Data Preparation Steps for Self-Service Business Intelligence 2212
Eugenie Lai, Yeye He, Surajit Chaudhuri

Time Series Motif Discovery: A Comprehensive Evaluation 2226
Valerio Guerrini, Thibaut Germain, Charles Truong, Laurent Oudre, Paul Boniol

EinDecomp: Decomposition of Declaratively-Specified Machine Learning and Numerical Computations for Parallel Execution 2240
Daniel Bourgeois, Zhimin Ding, Dimitrije Jankov, Jiehui Li, Mahmoud Sleem, Yuxin Tang, Jiawen Yao, Xinyu Yao, Chris Jermaine

Continuous Lifelong Conflict-Aware AGV Routing with Kinematic Constraints 2254
Ruizhong Wu, Mengxuan Zhang, Shuxin Wang, Kin Sun Chan, Yan Nei Law, Lei Li

Wolverine: Highly Efficient Monotonic Search Path Repair for Graph-based ANN Index Updates 2268
Dawei Liu, Bolong Zheng, Ziyang Yue, Fuhao Ruan, Xiaofang Zhou, Christian S. Jensen

Detecting Schema-Related Logic Bugs in Relational DBMSs via Equivalent Database Construction 2281
Jiansen Song, Wensheng Dou, Yingying Zheng, Yu Gao, Ziyu Cui, Wei Wang, Jun Wei

GraphSparseNet: a Novel Method for Large Scale Trafffic Flow Prediction 2295
Weiyang Kong, Wu Kaiqi, Sen Zhang, Yubao Liu

PVLDB ORGANIZATION AND REVIEW BOARD - Vol. 18

Editors in Chief of PVLDB
Themis Palpanas (University Paris Cite)
Nesime Tatbul (Intel Labs and MIT)

Associate Editors of PVLDB
Walid G. Aref (Purdue University)
Manos Athanassoulis (Boston University)
Carsten Binnig (Technical University of Darmstadt)
Spyros Blanas (Ohio State University)
Matthias Boehm (Technical University of Berlin)
Angela Bonifati (University of Lille)
K. Selcuk Candan (Arizona State University)
Lei Cao (University of Arizona)
Raul Castro Fernandez (University of Chicago)
Lei Chen (Hong Kong University of Science and Technology)
Shimin Chen (Chinese Academy of Sciences)
Yi Chen (New Jersey Institute of Technology)
Reynold Cheng (University of Hong Kong)
Alvin Cheung, University of California (Berkeley)
Sudipto Das (Amazon Web Services)
Niv Dayan (University of Toronto)
Antonis Deligiannakis (Technical University of Crete)
Jens Dittrich (Saarland University)
Xin Luna Dong (Meta)
Karima Echihabi (Mohammed VI Polytechnic University)
Alan Fekete (University of Sydney)
Avrilia Floratou (Microsoft)
Jana Giceva (Technical University of Munich)
Katja Hose (Technical University of Vienna)
H. V. Jagadish (University of Michigan)
Panos Kalnis (King Abdullah University of Science and Technology)
Georgia Koutrika (Athena Research Center)
Eric Lo (Chinese University of Hong Kong)
Nikos Mamoulis (University of Ioannina)
Stefan Manegold (CWI)
Ioana Manolescu (Inria and Polytechnic Institute of Paris)
Norman May (SAP SE)
Umar Farooq Minhas (Apple)
Fatemeh Nargesian (University of Rochester)
Beng Chin Ooi (National University of Singapore)
Fatma Ozcan (Google)
Tamer Ozsu (University of Waterloo)
Tilmann Rabl (Hasso Plattner Institute and University of Potsdam)
Mirek Riedewald (Northeastern University)
Jennie Rogers (Northwestern University)
Alkis Simitsis (Athena Research Center)
Letizia Tanca (Polytechnic University of Milan)
Nan Tang (Hong Kong University of Science and Technology (GZ))
Yuanyuan Tian (Microsoft)
Yongxin Tong (Beihang University)
Pinar Tozun (IT University of Copenhagen)
Yannis Velegrakis (Utrecht University)
Matthias Weidlich (Humboldt University of Berlin)
Steven E. Whang (Korea Advanced Institute of Science and Technology)
Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)

Publication Editors
Xiaoou Ding (Harbin Institute of Technology)
Subhadeep Sarkar (Brandeis University)
Giovanni Simonini (University of Modena and Reggio Emilia)

PVLDB Managing Editor
Jun Yang (Duke University)

PVLDB Advisory Board
Sourav S. Bhowmick (Nanyang Technological University)
Vanessa Braganholo (Universidade Federal Fluminense)
Lei Chen (Hong Kong University of Science and Technology)
Yanlei Diao (Ecole Polytechnique)
Xin Luna Dong (Meta)
Torsten Grust (University of Tuebingen)
Wolfgang Lehner (TU Dresden)
Alexandra Meliou (University of Massachusetts Amherst)
Felix Naumann (HPI)
Fatma Ozcan (Google)
Themis Palpanas (Universite Paris Cite)
Divesh Srivastava (AT&T Labs - Research)
Nesime Tatbul (Intel Labs and MIT)
Xiaokui Xiao (National University of Singapore)
Meihui Zhang (Beijing Institute of Technology)

Review Board

Ahmed S. Abdelhamid (Purdue University)
Ziawasch Abedjan (TU Berlin)
Ahmed Aly (Google)
Mohammad Javad Amiri (Stony Brook University)
Yael Amsterdamer (Bar-Ilan University)
Renzo Angles (Universidad de Talca)
Alexander Artikis (University of Piraeus)
Joy Arulraj (Georgia Tech)
Abolfazl Asudeh (University of Illinois Chicago)
Maurizio Atzori (University of Cagliari)
Nikolaus Augsten (University of Salzburg)
Zhifeng Bao (RMIT University)
Ilaria Bartolini (University of Bologna)
Johes Bater (Tufts University)
Lawrence Benson (HPI and University of Potsdam)
Sonia Bergamaschi (University of Modena and Reggio Emilia)
Anna Bernasconi (Politecnico di Milano)
Arnab Bhattacharya (IIT Kanpur)
Alexander Boehm (SAP SE)
Paul Boniol (Universite de Paris)
Renata Borovica-Gajic (University of Melbourne)
Panagiotis Bouros (Johannes Gutenberg University Mainz)
Vanessa Braganholo (Fluminense Federal University)
Matteo Brucato (Microsoft Research)
Michael J. Cahill (University of Sydney)
Diego Calvanese (Free University of Bozen Bolzano)
Jesus Camacho-Rodriguez (Microsoft)
Helena Caminal (Google)
Huiping Cao (New Mexico State University)
Yang Cao (University of Edinburgh)
Zhao Cao (Huawei Technologies)
Zhichao Cao (Arizona State University)
Matteo Ceccarello (University of Padova)
Chengliang Chai (Beijing Institute of Technology)
Yunpeng Chai (Renmin University of China)
Harry Kai-Ho Chan (The University of Sheffield)
Tsz Nam Chan (Shenzhen University)
Subarna Chatterjee (Harvard University)
Cindy Chen (University of Massachusetts Lowell)
Lu Chen (Zhejiang University)
Hong Cheng (The Chinese University of Hong Kong)
Rada Chirkova (NC State University)
Theodoros Chondrogiannis (University of Konstanz)
Shihabur Chowdhury (Apple)
George Christodoulou (TU Delft)
Periklis Chrysogelos (Oracle)
Gao Cong (Nanyang Technological Univesity)
Alex Conway (Cornell Tech)
Andrew Crotty (Northwestern University)
Bin Cui (Peking University)
Patrick Damme (TU Berlin)
Roshan Dathathri (Microsoft Research)
Jesse Davis (MongoDB)
Cagatay Demiralp (MIT)
Dong Deng (Rutgers University New Brunswick)
Laxman Dhulipala (University of Maryland, College Park)
Shimin Di (The Hong Kong University of Science and Technology)
Claudia Diamantini (Universita Politecnica delle Marche)
Anton Dignos (Free University of Bozen Bolzano)
Bailu Ding (Microsoft Research)
Bolin Ding (Alibaba Group)
Jialin Ding (Amazon Web Services)
Anh Dinh (Deakin University)
AnHai Doan (University of Wisconsin Madison)
Christos Doulkeridis (University of Pireaus)
Stefania Dumbrava (ENSIIE)
Ahmed Eldawy (University of California Riverside)
Mohamed Eltabakh (Qatar Foundation)
Venkatesh Emani (Microsoft)
Ju Fan (Renmin University of China)
Zhiwei Fan (Meta)
Yixiang Fang (The Chinese University of Hong Kong)
Anna Fariha (University of Utah)
Ziqiang Feng (Google)
Hakan Ferhatosmanoglu (University of Warwick and Amazon Web Services)
Elena Ferrari (University of Insubria)
Donatella Firmani (Sapienza University)
Peter M. Fischer (University of Augsburg)
George Fletcher (Eindhoven University of Technology)
Juliana Freire (New York University)
Sainyam Galhotra (Cornell University)
Johann Gamper (Free University of Bozen Bolzano)
Yunjun Gao (Zhejiang University)
Paolo Garza (Politecnico di Torino)
Chang Ge (University of Minnesota)
Tingjian Ge (University of Massachusetts Lowell)
Rainer Gemulla (Universitat Mannheim)
Nikos Giatrakos (Technical University of Crete)
Aristides Gionis (KTH Royal Institute of Technology)
Boris Glavic (Illinois Institute of Technology)
Lukasz Golab (University of Waterloo)
Jonathan Goldstein (Microsoft)
Sven Groppe (Universitat zu Lubeck)
Michael Grossniklaus (University of Konstanz)
Anja Gruenheid (Microsoft)
Le Gruenwald (The University of Oklahoma)
Vincenzo Gulisano (Chalmers University of Technology)
Rihan Hai (TU Delft)
Wook-Shin Han (POSTECH)
Mohamed S. Hassan (Oracle)
Oktie Hassanzadeh (IBM Research)
Wenjia He (University of Michigan)
Xi He (University of Waterloo)
Yeye He (Microsoft Research)
Meichun Hsu (Oracle)
Haibo Hu (The Hong Kong Polytechnic University)
Xiao Hu (University of Waterloo)
Qiang Huang (National University of Singapore)
Xin Huang (Hong Kong Baptist University)
Yan Huang (University of North Texas)
Zi Helen Huang (University of Queensland)
Madelon Hulsebos (University of California Berkeley)
Romain Ilbert (Huawei Paris Research Center)
Matteo Interlandi (Microsoft)

Ekaterini Ioanou (Tilburg University)
Gabriela Jacques-Silva (Facebook)
Fuad Jamour (Amazon Web Services)
Soren Kejser Jensen (Aalborg University)
Peiquan Jin (University of Science and Technology of China)
Alekh Jindal (SmartApps)
Hyungsoo Jung (Seoul National University)
Vasiliki Kalavri (Boston University)
Vana Kalogeraki (Athens University of Economics and Business)
Eser Kandogan (Megagon Labs)
Daniel Kang (UIUC)
Zoi Kaoudi (IT University of Copenhagen)
Pinar Karagoz (Middle East Technical University (METU))
Bojan Karlas (Harvard University)
Asterios Katsifodimos (TU Delft)
Oliver A. Kennedy (University at Buffalo SUNY)
Arijit Khan (Aalborg University)
Guy Khazma (University of Toronto)
Haridimos Kondylakis (FORTH-ICS)
Arnd Christian Konig (Microsoft)
Chrysanthi Kosyfaki (The University of Hong Kong)
Nick Koudas (University of Toronto)
Paraschos Koutris (University of Wisconsin Madison)
Mayuresh Kunjir (Amazon Web Services)
Alexandros Labrinidis (University of Pittsburgh)
Wolfgang Lehner (TU Dresden)
Chuan Lei (Amazon Web Services)
Viktor Leis (TU Munich)
Alberto Lerner (University of Fribourg)
Ulf Leser (Humboldt-Universitat zu Berlin)
Guoliang Li (Tsinghua University)
Jia Li (The Hong Kong University of Science and Technology (GZ))
Jianxin Li (Deakin University)
Tian Li (Carnegie Mellon University)
Tianyu Li (MIT)
Yinan Li (Microsoft Research)
Yuchen Li (Singapore Management University)
Xiang Lian (Kent State University)
Shen Liang (Universite Paris Cite)
Michele Linardi (CYU)
Matteo Lissandrini (University of Verona)
Chunwei Liu (MIT)
Jinfei Liu (Zhejiang University)
Xueli Liu (Tianjin University)
Cheng Long (Nanyang Technological University)
Baotong Lu (Microsoft Research)
Jiaheng Lu (University of Helsinki)
Siqiang Luo (Nanyang Technological University)
Yuyu Luo (The Hong Kong University of Science and Technology (GZ))
Manisha Luthra (TU Darmstadt)
Joana M. F. da Trindade (MIT)
Chenhao Ma (The Chinese University of Hong Kong)
Lin Ma (University of Michigan)
Amr Magdy (University of California Riverside)
Ahmed Mahmood (Google)
Sujaya Maiyya (University of Waterloo)
Neha Makhija (Northeastern University)
Silviu Maniu (Universite Grenoble Alpes)
Essam Mansour (Concordia University)
Ryan Marcus (University of Pennsylvania)
Amelie Marian (Rutgers University)
Davide Martinenghi (Politecnico di Milano)
Venkata Vamsikrishna Meduri (IBM Research - Almaden)
Sharad Mehrotra (University of California Irvine)
Alexandra Meliou (University of Massachusetts Amherst)
Paolo Merialdo (Universita degli Studi Roma Tre)
Amine Mhedhbi (Polytechnique Montreal)
Xiaoye Miao (Zhejiang University)
Sebastian Michel (RPTU Kaiserslautern Landau)
Katsiaryna Mirylenka (IBM Research Zurich)
Madhulika Mohanty (Inria Saclay)
Mohamed Mokbel (University of Minnesota Twin Cities)
Mirella M. Moro (Universidade Federal de Minas Gerais)
Davide Mottin (Aarhus University)
Kyriakos Mouratidis (Singapore Management University)
Ingo Müller (Google)
Balakrishnan Narayanaswamy (Amazon)
Mario Nascimento (Northeastern University)
Parimarjan Negi (MIT)
Quoc Viet Hung Nguyen (Griffith University)
Milos Nikolic (University of Edinburgh)
Matthaios Olma (MongoDB)
Prashant Pandey (University of Utah)
George Papadakis (University of Athens)
Dimitris Papadias (The Hong Kong University of Science and Technology)
Odysseas Papapetrou (TU Eindhoven)
John Paparrizos (The Ohio State University)
George Papastefanatos (ATHENA Research Center)
Stefano Paraboschi (Universita degli Studi di Bergamo)
Aditya Parameswaran (University of California Berkeley)
Yongjoo Park (UIUC)
Eliana Pastor (Politecnico di Torino)
Jignesh Patel (Carnegie Mellon University)
Marco Patella (University of Bologna)
Torben Bach Pedersen (Aalborg University)
Botao Peng (Chinese Academy of Sciences)
Peng Peng (Hunan University)
Matthew J. Perron (MIT)
Ilia Petrov (Reutlingen University)
Holger Pirk (Imperial College)
Stefan Plantikow (Neo4j)
Orestis Polychroniou (Amazon)
Danica Porobic (Oracle)
Abdulhakim Qahtan (Utrecht University)
Abdul Quamar (Google)
Weixiong Rao (Tongji University)
Berthold Reinwald (IBM Research Almaden)
El Kindi Rezig (MIT)
Daniel Ritter (SAP)
Oscar Romero (Universitat Politecnica de Catalunya)
Kexin Rong (Georgia Institute of Technology)
Abhishek Roy (Snowflake)

Florin Rusu (University of California Merced)
Sourav S. Bhowmick (Nanyang Technological University)
Ibrahim Sabek (University of Southern California)
Mohammad Sadoghi (University of California Davis)
Semih Salihoglu (University of Waterloo)
Maria Luisa Sapino (University of Torino)
Subhadeep Sarkar (Brandeis University)
Kai-Uwe Sattler (TU Ilmenau)
Patrick Schafer (Humboldt-Universitat zu Berlin)
Felix M. Schuhknecht (Johannes Gutenberg University Mainz)
Maximilian E. Schule (University of Bamberg)
Malte Schwarzkopf (Brown University)
Rathijit Sen (Microsoft)
Jiwon Seo (Seoul National University)
Juan Sequeda (data.world)
Marco Serafini (University of Massachusetts Amherst)
Amir Shaikhha (University of Edinburgh)
Shantanu Sharma (New Jersey Institute of Technology)
Yanyan Shen (Shanghai Jiao Tong University)
Jieming Shi (The Hong Kong Polytechnic University)
Roee Shraga (WPI)
Tarique Siddiqui (Microsoft Research)
Giovanni Simonini (University of Modena and Reggio Emilia)
Utku Sirin (Harvard University)
Spiros Skiadopoulos (University of the Peloponnese)
Dimitrios Skoutas (Athena Research Center)
Shaoxu Song (Tsinghua University)
Divesh Srivastava (AT&T Chief Data Office)
Chrysoula Stathakopoulou (Chainlink Labs)
Kostas Stefanidis (Tampere University)
Kurt Stockinger (ZHAW Zurich University of Applied Sciences)
Uta Storl (University of Hagen)
Shixuan Sun (Shanghai Jiao Tong University)
Ki Hyun Tae (KAIST)
Dixin Tang (University of Texas Austin)
Jing Tang (The Hong Kong University of Science and Technology (GZ))
Mingjie Tang (Sichuan University)
Bo Tang (Southern University of Science and Technology)
Egemen Tanin (University of Melbourne)
Ernest Teniente (Universitat Politecnica de Catalunya)
Arash Termehchy (Oregon State University)
Jens Teubner (TU Dortmund)
Riccardo Torlone (Roma Tre University)
Goce Trajcevski (Iowa State University)
Immanuel Trummer (Cornell University)
Eleni Tzirita Zacharatou (IT University of Copenhagen)
Katerina Tzompanaki (CY Cergy Paris University)
Leong Hou U (University of Macau)
Alexander van Renen (UTN)
Genoveva Vargas-Solar (CNRS LIRIS)
Nalini Venkatasubramanian (University of California Irvine)
Hannes Voigt (Neo4j)
Hongzhi Wang (Harbin Institute of Technology)
Ning Wang (Beijing Jiaotong University)
Qitong Wang (Universite Paris Cite)
Sibo Wang (The Chinese University of Hong Kong)
Tianzheng Wang (Simon Fraser University)
Yifan Wang (University of Florida)
Sai Wu (Zhejiang University)
Yinghui Wu (Case Western Reserve University)
Yuncheng Wu (Renmin University of China)
Xiaokui Xiao (National University of Singapore)
Jianliang Xu (Hong Kong Baptist University)
Jianqiu Xu (Nanjing University of Aeronautics and Astronautics)
Nikolay Yakovets (TU Eindhoven)
Xiao Yan (Centre for Perceptual and Interactive Intelligence (CPII))
Hongzhi Yin (The University of Queensland)
Man Lung Yiu (The Hong Kong Polytechnic University)
Brit Youngmann (Technion)
Jeffrey Xu Yu (The Chinese University of Hong Kong)
Xiaohui Yu (York University)
Yi Yu (NII)
Ye Yuan (Beijing Institute of Technology)
Cong Yue (National University of Singapore)
Demetrios Zeinalipour-Yazti (University of Cyprus)
Yuxiang Zeng (Beihang University)
Steffen Zeuch (TU Berlin)
Chao Zhang (University of Waterloo)
Chen Zhang (The Hong Kong Polytechnic University)
Huanchen Zhang (Tsinghua University)
Meihui Zhang (Beijing Institute of Technology)
Minjia Zhang (Microsoft AI and Research)
Qizhen Zhang (University of Toronto)
Xiaofei Zhang (University of Memphis)
Yanfeng Zhang (Northeastern University)
Bo Zhao (Aalto University)
Zhuoyue Zhao (University at Buffalo)
Bolong Zheng (Huazhong University of Science and Technology)
Kaiping Zheng (National University of Singapore)
Jingren Zhou (Alibaba Group)
Xuan Zhou (East China Normal University)
Yongluan Zhou (University of Copenhagen)
Yiwen Zhu (Microsoft)
Jia Zou (Arizona State University)
Lei Zou (Peking University)
Kostas Zoumpatianos (Snowflake)
Andreas Zufle (Emory University)

LETTER FROM THE EDITORS IN CHIEF

We are pleased to present the seventh (March 2025) issue of PVLDB Volume 18, which contains 23 papers from three categories: Scalable Data Science (2), Experiment, Analysis & Benchmark (1), and Regular Research (20).

These papers cover a diverse collection of primary subject areas, including: Data Mining and Analytics (4); Data Privacy and Security (3); Graph and Network Data (3); Machine Learning, AI, and Databases (3); Specialized and Domain-Specific Data Management (3); Database Engines (2); Information Integration and Data Quality (2); Distributed Database Systems (1); Provenance and Workflows (1); and User Interfaces (1).

The Scalable Data Science (SDS) paper category was first launched as part of PVLDB Volume 14, and since then, has seen growing interest from our database community. These papers focus on the design, implementation, experience, or evaluation of solutions and systems for practical data science and data engineering tasks, including data management, data engineering, data analytics, data visualization, data quality, data integration, data mining, and machine learning on large-scale data. Distinct from the Regular Research papers, papers in this category do not necessarily propose new algorithms or models, but emphasize solutions that either solve or advance the understanding of issues related to data science technologies in the real world.

In Volume 18, we reviewed a total of 81 SDS submissions to the PVLDB research track. In the first six issues published so far, there appeared 10 accepted SDS papers. In this seventh issue, we have 2 new ones: one presenting a middleware approach to data science query federation, and another proposing a scalable model pre-training approach for spatio-temporal predictions.

Overall, many of these SDS papers describe end-to-end systems for data science, machine learning, search, and prescriptive analytics; or alternatively, share specific primitives and algorithms for data-centric AI. The dedicated SDS paper category provides an important forum for such papers that describe new system design points, which compose or orchestrate existing methods in scalable and effective manners. Interestingly, many of the accepted papers also exhibit a good share of collaborations across industry and academia.

We are grateful to our board of associate editors and reviewers as well as proceedings chairs for their dedicated service that made this issue possible, with special thanks to our associate editors overseeing the SDS category.

Matthias Boehm, Reynod Cheng, Xin Luna Dong
Associate Editors for the Scalable Data Science (SDS) Paper Category

Themis Palpanas and Nesime Tatbul
Editors-in-Chief of PVLDB Vol. 18
Program Chairs for VLDB 2025