



Proceedings of the VLDB Endowment

Volume 19, No. 9 – May 2026

Editors in Chief:

Yanlei Diao and Xiaokui Xiao

Associate Editors:

Alkis Polyzotis, Alkis Simitsis, Andrew Pavlo, Angela Bonifati, Anh Dinh, Antonis Deligiannakis, Ashraf Aboulnaga, Badrish Chandramouli, Bin Cui, Bogdan Cautis, Ce Zhang, Elena Ferrari, Evaggelia Pitoura, Fatemeh Nargesian, Guoliang Li, Hong Cheng, Immanuel Trummer, Jana Giceva, Jennie Rogers, Jian Pei, Jianliang Xu, Karima Echihabi, Katja Hose, Khuzaima Daudjee, Kunsoo Park, Kyuseok Shim, Laks Lakshmanan, Lei Cao, Lei Chen, Li Xiong, Mauro Sozio, Meihui Zhang, Melanie Herschel, Michael Böhlen, Nan Tang, Niv Dayan, Norman May, Raul Castro Fernandez, Raymond Chi-Wing Wong, Reynold Cheng, Ruben Mayer, Semih Salihoglu, Shimin Chen, Sibor Wang, Sourav S Bhowmick, Spyros Blanas, Steven Whang, Sudipto Das, Umar Farooq Minhas, Wei Wang, Wook-Shin Han, Yannis Velegrakis, Yanyan Shen, Yi Chen, Yuncheng Wu, Zhifeng Bao, Zi Huang

Publication Editors:

Andrea Mauri, Giovanni Simonini, Siqiang Luo, Stefan Halfpap, Subhadeep Sarkar

PVLDB – Proceedings of the VLDB Endowment

Volume 19, No. 9, May 2026.

All papers published in this issue will be presented at the 52nd International Conference on Very Large Data Bases, Boston, MA, USA, 2026.

Copyright 2026 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 19, Number 9, May 2026

Pages i – xi and 1854 - 2535

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. 19	vi

Research Papers

Secure Join Operations in Multi-Identifier Databases: Performance and Practicality	1854
<i>Wen-Jie Lu, Yongchuan Niu, Yongjun Zhao, Wei Dai, Donghang Lu, Li Wang, Qiang Yan</i>	
A Resource-centric Analysis and Optimization of NoSQL Workloads using Distressed Resource Volume Metric.....	1867
<i>Gunika Verma, Aashutosh A V, Pooja Srinivas, Yogesh Simmhan, Ayush Choure, Harshit Shah, Mayukh Das, Prashant Sasatte, Chetan Bansal, Abhijit Pai, Suraj Dixit, Achint Agrawal</i>	
Efficient Banzhaf-Based Data Valuation for k-Nearest Neighbors Classification	1880
<i>Guangyi Zhang, Lutz Oettershagen, Lixu Wang, Aristides Gionis</i>	
Mil: Cost-guided Minimum Makespan Scheduling for Applications of Multiple LLMs.....	1893
<i>Jingzhi Fang, Yanyan Shen, Yue Wang, Lei Chen</i>	
Dynamic read & write optimization with TurtleKV	1907
<i>Tony Astolfi, Vidya Silai, Darby Huye, Lan Liu, Raja R. Sambasivan, Johes Bater</i>	
Compass: SLO-aware Query Planner for Compound AI Serving at Scale.....	1921
<i>Banruo Liu, Wei-Yu Lin, Minghao Fang, Yihan Jiang, Fan Lai</i>	
KDSelector: A Framework of Knowledge-Enhanced and Data-Efficient Selector Learning for Anomaly Detection Model Selection in Time Series.....	1935
<i>Zhiyu Liang, Dongrui Cai, Chenyuan Zhang, Zheng Liang, Chen Liang, Shi Qiu, Jin Wang, Hongzhi Wang</i>	
Multimodal Knowledge Graph Completion via Relation-Aware Negative Sampling with Diffusion-based Interpolation	1949
<i>Qian Ma, Linfei Dai, Zhongming Yao, Yu Gu, Tianyi Li, Christian S. Jensen, Ge Yu</i>	
RT-RkNN: Reverse k Nearest Neighbor Queries as a Graphics Ray Casting Problem	1963
<i>Zhengyang Bai, Peng Chen, Mohamed Wahib</i>	
Swan: Hybrid MVCC Management for Efficient Transaction Processing in LSM-Tree-Based Key-Value Stores	1977
<i>Yang Guo, Jin Xue, Zili Shao</i>	
Kirin: Efficient In-Storage Learned Compaction for LSM-Trees via System-Algorithm Co-Design.....	1991
<i>Guifeng Wang, Shengan Zheng, Penghao Sun, Jin Pu, Kaijiang Deng, Bowen Zhang, Weihang Kong, Cong Zhou, Yifan Hua, Linpeng Huang</i>	
LLMs as Stratification Signals for KG Accuracy Evaluation	2005
<i>Stefano Marchesin, Matteo Ceccarelli, Gianmaria Silvello</i>	
A Comparative Evaluation of Schema Subsetting for LLM-based NL-to-SQL over Large-Schema Databases	2019

Kyle Luoma, Arun Kumar

IncreQueryFusion: On-demand Data Fusion Framework in Dynamic Data Lakes	2032
<i>Wenhao Liu, Sai Wu, Xiu Tang, Yitong Zhang, Dong Peng, Guolong Huang, Gang Chen</i>	
One Join Order Does Not Fit All: Reducing Intermediate Results with Per-Split Query Plans	2045
<i>Yujun He, Hangdong Zhao, Simon Frisk, Yifei Yang, Kevin Kristensen, Paraschos Koutris, Xiangyao Yu</i>	
Harmonizing Efficiency and Accuracy in Filtered Vector Search	2059
<i>Zixiang Zhou, Xuhao Chen</i>	
Counting HyperGraphlets via Color Coding: a Quadratic Barrier and How to Break It.....	2073
<i>Marco Bressan, Stefano Clemente, Giacomo Fumagalli</i>	
Discovering Approximate Denial Constraints in Large Databases.....	2086
<i>Albert Martin, Eduardo C. De Almeida, Oscar Romero, Anna Queralt</i>	
Human-Centered Exploration of Table Unionability	2099
<i>Nina Klimenkova, Sreeram Marimuthu, Roei Shraga</i>	
Redbench: Workload Synthesis From Cloud Traces	2113
<i>Johannes Wehrstein, Roman Heinrich, Mihail Stoian, Skander Krid, Martin Stemmer, Andreas Kipf, Carsten Binnig, Muhammad El-Hindi</i>	
HarborMaster: Rollback Detection for Trusted Distributed Computing	2126
<i>Shubham Mishra, Alexander Thomas, Nurzhan Abdrassilov, Kaiyuan Chen, Natacha Crooks, John Kubiatawicz</i>	
Auditing for Demographic Bias in Opaque Rankings.....	2140
<i>Antonio Ferrara, Carlo Abrate, Fabio Vitale, Francesco Bonchi</i>	
Efficient Cooperation-Aware Key and Value Management for LLM Inference	2154
<i>Qiheng Sun, Hongwei Zhang, Junxu Liu, Haocheng Xia, Jinfei Liu, Kui Ren, Haibo Hu</i>	
Succinct and Fast Tiny Pointer Hash Tables	2168
<i>Xilin Tang, Yuqi Mai, William Kuszmaul, Alex Conway</i>	
Aggregating maximal cliques in real-world graphs.....	2183
<i>Noga Alon, Sabyasachi Basu, Shweta Jain, Haim Kaplan, Jakub Łącki, Blair D. Sullivan</i>	
PRISM: A Training System to Unlock the Potential of Temporal Graph Learning Through Staleness Avoidance	2196
<i>Md Ashraful Islam, Hojae Son, Suhaas Kiran Doddagaddavalli Gangadharaiah, Marco Serafini</i>	
SafeQL: Search-based Refinement for Safe and Efficient LLM-based Text-to-SQL	2210
<i>Geonho Lee, Min-Soo Kim</i>	
QA-GraphRAG: Query-Adaptive Plug-and-Play Retrieval Integration for Graph-based Retrieval-Augmented Generation	2224
<i>Zeang Sheng, Ruihong Sun, Jiahao Xu, Hanmei Luo, Peng Chen, Wentao Zhang, Bin Cui</i>	
Testing Graph Databases via Transformations Between Fixed-Length and Variable-Length Queries.....	2234
<i>Jinxin Gui, Yuanhong Lan, Longlong Lu, Yifei Lu, Minxue Pan</i>	
Efficient, Scalable, and Fair Locking on Disaggregated Memory with Decentralized Coordination.....	2248
<i>Hanze Zhang, Ke Cheng, Rong Chen, Xingda Wei, Haibo Chen</i>	

MGI: A Communication Framework for Data Processing in Massive GPU Infrastructures.....	2262
<i>Di Wu, Hongshi Tan, Hanzhang Yang, Bingsheng He, Qizhen Zhang</i>	
MDS-FSM: Coverage-Based Frequent Subgraph Mining in Single Graphs	2276
<i>Xiaozhen Guo, Xueli Liu, Bowen Dong, Li Wan, Jiake Ge, Shuai Ma</i>	
Continuous Query for Top-K Maximal Sum Intervals over Streaming Data.....	2289
<i>Zhongshuai Zhang, Xiaochun Yang, Baihua Zheng, Rui Zhu, Haomin Li, Bin Wang</i>	
Subgraph Enumeration: Beyond Tree Decomposition	2303
<i>Qiyang Li, Jeffrey Xu Yu, Zongyan He</i>	
High-Performance DBMSs with io_uring: When and How to Use It.....	2317
<i>Matthias Jasny, Muhammad El-Hindi, Tobias Ziegler, Viktor Leis, Carsten Binnig</i>	
CREST: Approximate k-Clique Counting in Real-World Networks via Refinement of Star-Based Sample Space	2331
<i>Yehyun Nam, Jihoon Jang, Kunsoo Park, Joong Chae Na, Hyunjoon Kim</i>	
FutureLight: An Efficient Future Traffic Data-Driven Reinforcement Learning Framework for Traffic Signal Controls.....	2344
<i>Zizhuo Xu, Haolun Ma, Lei Li, Zhiyuan Wang, Yunjie Huang, Xiaofang Zhou</i>	
BookRAG: A Hierarchical Structure-aware Index-based Approach for Retrieval-Augmented Generation on Complex Documents.....	2358
<i>Shu Wang, Yingli Zhou, Yixiang Fang</i>	
Replacing Multi-Step Assembly of Data Preparation Pipelines with One-Step LLM Pipeline Generation for Table QA	2372
<i>Fengyu Li, Junhao Zhu, Kaishi Song, Lu Chen, Zhongming Yao, Tianyi Li, Christian S. Jensen</i>	
Storage-Centric Relation Design via High-Quality Approximate Functional Dependencies	2385
<i>Rui Ding, Xiaochun Yang, Bin Wang, Quanqing Xu, Chuanhui Yang</i>	
Unstructured Data Analysis using LLMs: A Comprehensive Benchmark	2398
<i>Qiyang Deng, Jianhui Li, Chengliang Chai, Ye Yuan, Jinqi Liu, Junzhi She, Kaisen Jin, Zhaoze Sun, Yuhao Deng, Jia Yuan, Yuping Wang, Xu Zhou, Guoren Wang, Lei Cao</i>	
Measuring Database Unfairness via Dependency Quantification Under Differential Privacy	2411
<i>Mariia Vologdin, Yuchao Tao, Amir Gilad</i>	
STEM2: A Fast and Space-efficient Data Structure for Exact Multi-Set Membership Query	2426
<i>Yannian Niu, Song Han, Minmei Wang</i>	
Revisiting the Maximum Defective Clique Problem: Faster Branching and a Tighter Upper Bound	2439
<i>Kewu Yang, Kaiqiang Yu, Shengxin Liu, Zhaoquan Gu</i>	
NeurIDA: Dynamic Modeling for Effective In-Database Analytics	2452
<i>Lingze Zeng, Shaofeng Cai, Naili Xing, Jiaqi Zhu, Gang Chen, Peng Lu, Jian Pei, Beng Chin Ooi</i>	
Finding Non-Redundant Simpson's Paradox in Multidimensional Data.....	2466
<i>Yi Yang, Jian Pei, Jun Yang, Jichun Xie</i>	

MGRAG: Semantic Subgraph Matching and Graph-Aware Caching for Multimodal Retrieval-Augmented Generation	2480
<i>Yubo Wang, Haoyang Li, Lei Chen</i>	
PystachIO: Efficient Distributed GPU Query Processing with PyTorch over Fast Networks & Fast Storage	2494
<i>Jigao Luo, Nils Boeschen, Muhammad El-Hindi, Carsten Binnig</i>	
BaCon: Efficient Batch Processing of Counting Queries	2508
<i>Yuxi Liu, Xiao Hu, Pankaj K. Agarwal, Jun Yang</i>	
Document-to-Database: Extraction Meets Relational Semantics	2522
<i>Zhengxuan Zhang, Zhuowen Liang, Jiazhao Chen, Haixun Wang, Nan Tang</i>	

PVLDB ORGANIZATION AND REVIEW BOARD - Vol. 19

Editors in Chief of PVLDB

Yanlei Diao (Ecole Polytechnique and Amazon Web Services)
Xiaokui Xiao (National University of Singapore)

Associate Editors of PVLDB

Alkis Polyzotis (Databricks)
Alkis Simitsis (Athena Research Center)
Andrew Pavlo (Carnegie Mellon University)
Angela Bonifati (Univ. of Lyon)
Anh Dinh (Deakin University)
Antonis Deligiannakis (Technical University of Crete)
Ashraf Aboulmaga (University of Texas at Arlington)
Badrish Chandramouli (Microsoft Research)
Bin Cui (Peking University)
Bogdan Cautis (University of Paris-Saclay)
Ce Zhang (University of Chicago)
Elena Ferrari (University of Insubria, Varese)
Evaggelia Pitoura (Univ. of Ioannina)
Fatemeh Nargesian (University of Rochester)
Guoliang Li (Tsinghua University)
Hong Cheng (Chinese University of Hong Kong)
Immanuel Trummer (Cornell University)
Jana Giceva (TU Munich)
Jennie Rogers (Northwestern University)
Jian Pei (Duke University)
Jianliang Xu (Hong Kong Baptist University)
Karima Echiabi (Mohammed VI Polytechnic University)
Katja Hose (TU Wien)
Khuzaima Daudjee (University of Waterloo)
Kunsoo Park (Seoul National University)
Kyuseok Shim (Seoul National University)
Laks Lakshmanan (The University of British Columbia)
Lei Cao (University of Arizona)
Lei Chen (Hong Kong University of Science and Technology)
Li Xiong (Emory University)
Mauro Sozio (Luiss University of Rome)
Meihui Zhang (Beijing Institute of Technology)
Melanie Herschel (Nanyang Technological University)
Michael Böhlen (University of Zurich)
Nan Tang (HKUST (GZ))
Niv Dayan (University of Toronto)
Norman May (SAP SE)
Raul Castro Fernandez (The University of Chicago)
Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)
Reynold Cheng (The University of Hong Kong, China)
Ruben Mayer (University of Bayreuth)
Semih Salihoglu (University of Waterloo)
Shimin Chen (Chinese Academy of Sciences)
Sibo Wang (The Chinese University of Hong Kong)
Sourav S Bhowmick (Nanyang Technological University)
Spyros Blanas (The Ohio State University)
Steven Whang (KAIST)

Sudipto Das (Amazon Web Services)
Umar Farooq Minhas (Apple)
Wei Wang (ByteDance)
Wook-Shin Han (POSTECH)
Yannis Velegrakis (Utrecht University)
Yanyan Shen (Shanghai Jiao Tong University)
Yi Chen (NJIT)
Yuncheng Wu (Renmin University of China)
Zhifeng Bao (The University of Queensland)
Zi Huang (University of Queensland)

Publication Editors

Andrea Mauri (Lyon 1 University)
Giovanni Simonini (University of Modena and Reggio Emilia)
Siqiang Luo (Nanyang Technological University)
Stefan Halfpap (TU Berlin)
Subhadeep Sarkar (Brandeis University)

PVLDB Managing Editor

Jun Yang (Duke University)

PVLDB Advisory Board

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

Review Board

Abhishek Roy (Microsoft)
Ahmed Aly (Google)
Ahmed Eldawy (University of California, Riverside)
Ahmed Mahmood (Google)
Alberto Lerner (Computing Flows GmbH)
Alberto Sonnino (MystenLabs & University College London)
Alexander Artikis (University of Piraeus)
Alexander Böhm (SAP SE)
Alexander Krause (TU Dresden)
Amarnadh Sai Eluri (Google)
Amelie Marian (Rutgers University)
Amine Mhedhbi (Polytechnique Montréal)
Amir Shaikhha (University of Edinburgh)
Amr Magdy (University of California Riverside)
Andrea Mauri (Lyon 1 University)
Andreas Züfle (Emory University)
Antonio Boffa (EPFL)
Arash Termehchy (Oregon State University)
Arijit Khan (Bowling Green State University)
Aristides Gionis (KTH Royal Institute of Technology)
Arkaprava Saha (University of Tsukuba)
Arnab Phani (TU Berlin)
Arnd Christian König (Microsoft)
Avinash Kumar (Google)
Baihua Zheng (Singapore Management University)
Berthold Reinwald (IBM Research - Almaden)
Bin Yang (Aalborg University)
Bo Tang (Southern University of Science and Technology)
Bo Zhao (Aalto University)
Bobbi W Yogatama (NVIDIA)
Bojan Karlaš (Harvard University)
Boris Glavic (University of Illinois Chicago)
Brian Kroth (Microsoft)
Brit Youngmann (Technion)
Calisto Zuzarte (IBM)
Carlo Curino (Microsoft)
Chang Ge (University of Minnesota)
Chao Zhang (University of Waterloo)
Chen Zhang (The Hong Kong Polytechnic University)
Cheng Long (Nanyang Technological University)
Chenghao Lyu (Amazon)
Chengkai Li (The University of Texas at Arlington)
Chengliang Chai (Beijing Institute of Technology)
Chenhao Ma (The Chinese University of Hong Kong, Shenzhen)
Christan Grant (University of Florida)
Christos Doukeridis (University of Piraeus)
Christos Koutras (New York University)
Chrysanthi Kosyfaki (The University of Hong Kong)
Chuan Lei (Oracle)
Chunwei Liu (MIT CSAIL)
Cindy Chen (University of Massachusetts Lowell)
Cong Yue (National University of Singapore)
Daniel Kang (UIUC)
Daniel Ritter (SAP)
Davood Rafiei (University of Alberta)
Demetris Zeinalipour (University of Cyprus)
Dimitrios Skoutas (Athena Research Center)

Dixin Tang (The University of Texas, Austin)
Domagoj Vrgoč (PUC)
Dominik Durner (CedarDB)
Donatella Firmani (Sapienza University)
Donato Tiano (Università degli Studi di Modena e Reggio Emilia)
Dong Deng (Alibaba Cloud Computing)
Dong Xie (Penn State University)
Dumitrel Loghin (OKX)
Eduard Dragut (Temple Univ.)
Egemen Tanin (University of Melbourne)
El Kindi Rezig (University of Utah)
Eleni Tziritza Zacharitou (Hasso Plattner Institute (HPI))
Eliana Pastor (Politecnico di Torino)
Elisa Bertino (Purdue University)
Enzo Veltri (Università della Basilicata)
Ergute Bao (Mohamed bin Zayed University of Artificial Intelligence)
Eser Kandogan (Megagon Labs)
Essam Mansour (Concordia University)
Farnoush Banaei-Kashani (University of Colorado Denver)
Felix M Schuhknecht (Johannes Gutenberg University Mainz)
Florian Hahn (University of Twente)
Florin Rusu (UC Merced)
Francesco Gullo (University of L'Aquila)
Fuad Jamour (Amazon Web Services)
Gao Cong (Nanyang Technological University)
Garima Gaur (Inria Saclay, Ecole Polytechnique)
Gengrui Zhang (Concordia University)
George Christodoulou (TU Delft)
George Papadakis (University of Athens)
George Papastefanatos (ATHENA Research Center)
Georgios Siachamis (Inria)
Georgios R Theodorakis (Nvidia)
Giorgio Vinciguerra (Università di Pisa)
Giuseppina Andresini (University of Bari Aldo Moro)
Goce Trajcevski (Iowa State University)
Grigorios Loukides (King's College London)
Guido Moerkotte (University of Mannheim)
Guillaume Lachaud (Ecole Polytechnique)
Haibo Hu (Hong Kong Polytechnic University)
Haipeng Dai (Nanjing University)
Hannes Voigt (Neo4j)
Hantian Zhang (Google)
Hanzhi Wang (University of Copenhagen)
Haralampos Gavrilidis (BIFOLD & Technische Universität Berlin)
Haridimos Kondylakis (FORTH-ICS & Computer Science Department, University of Crete)
Harry Kai-Ho Chan (The University of Sheffield)
Hazar Harmouch (University of Amsterdam)
Herald Kllapi (Google Inc)
Hongzhi Wang (Harbin Institute of Technology)
Hua Lu (Aalborg University)
Huanchen Zhang (Tsinghua University)
Huiping Chen (University of Birmingham)
Hyungsoo Jung (Seoul National University)
Hyunjoon Kim (Hanyang University)
Ibrahim Sabek (University of Southern California)

Igor Zablotchi (Mysten Labs)
 Ilia Petrov (Reutlingen University)
 Ilias Azizi (Mohammed VI Polytechnic University)
 Ilie Sarpe (KTH Royal Institute of Technology)
 Indrakshi Ray (Colorado State University)
 Ioannis Demertzis (UCSC)
 Jaeyoung Do (Seoul National University)
 Jan Hidders (Birkbeck, University of London)
 Jeeta Ann Chacko (Technical University of Munich)
 Jeffrey Xu Yu (The Hong Kong University of Science and Technology (Guangzhou))
 Jelle Hellings (McMaster University)
 Jesús Camacho-Rodríguez (Microsoft)
 Jia Li (Hong Kong University of Science and Technology)
 Jia Zou (Arizona State University)
 Jialin Ding (Princeton University)
 Jianguo Wang (Purdue University)
 Jiannan Wang (Tsinghua University)
 Jianxin Li (Edith Cowan University)
 Jiawei Jiang (Wuhan University)
 Jieming Shi (The Hong Kong Polytechnic University)
 Jignesh Patel (Carnegie Mellon University)
 Jin Wang (Arizona State University)
 Jiwon Seo (Seoul National University)
 Johannes Pietrzyk (TU Dresden)
 Johes Bater (Tufts University)
 John Paparrizos (The Ohio State University, USA & Aristotle University of Thessaloniki, Greece)
 Jonathan Fürst (ZHAW Zurich University of Applied Sciences)
 Jonghyeok Park (Korea University)
 Joseph Near (University of Vermont)
 Ju Fan (Renmin University of China)
 Junghoon Kim (UNIST)
 Junhao Gan (University of Melbourne)
 K. Selçuk Candan (Arizona State University)
 Kaisong Huang (University of Calgary)
 Kai-Uwe Sattler (TU Ilmenau)
 Kanchan Chowdhury (Marquette University)
 Karim Benouaret (Univ Lyon)
 Katsiaryna Milylenka (Zalando SE)
 Kaustubh Beedkar (Indian Institute of Technology Delhi)
 Keke HUANG (The University of British Columbia)
 Kenneth A Ross (Columbia University)
 KI HYUN TAE (Samsung Research)
 Kostas Stefanidis (Tampere University)
 Kurt Stockinger (Zurich University of Applied Sciences)
 Kwanghyun Park (Yonsei University)
 Kyoungmin Kim (EPFL)
 Larissa Capobianco Shimomura (Hasselt University)
 Le Gruenwald (The University of Oklahoma)
 Lefteris Kokoris Kogias (MystenLabs)
 Lei Zou (Peking University)
 Leong Hou U (University of Macau)
 Leopoldo Bertossi (Carleton University)
 Liang Liang (EPFL)
 Lingyang Chu (McMaster University)
 Loredana Caruccio (University of Salerno)
 Lu Chen (Zhejiang University)
 Luca Gagliardelli (eCampus University)
 Lucas Braun (Oracle Labs)
 Lukasz Golab (University of Waterloo)
 Madhulika Mohanty (Inria & Institut Polytechnique de Paris)
 Magnus Mueller (AWS)
 Mahdi Esmailoghli (HU Berlin)
 Makoto Onizuka (Osaka University)
 Man Lung Yiu (Hong Kong Polytechnic University)
 Manuel Rigger (National University of Singapore)
 Marco Patella (University of Bologna)
 Markos Markakis (Massachusetts Institute of Technology)
 Martin Prammer (Carnegie Mellon University)
 Matteo Ceccarello (University of Padova)
 Matteo Interlandi (Microsoft)
 Matthaïos Olma (MongoDB)
 Matthew J Perron (Microsoft)
 Matthias Boehm (Technische Universität Berlin)
 Matthias Weidlich (Humboldt-Universität zu Berlin)
 Maximilian E Schüle (University of Bamberg)
 Meichun Hsu (Oracle America)
 Miao Qiao (The University of Auckland)
 Miao Yu (Meta)
 Michael J Cahill (University of Sydney)
 Michael Grossniklaus (University of Konstanz)
 Michael Gubanov (Florida State University)
 Michael J Mior (Rochester Institute of Technology)
 Mohamed Eltabakh (Qatar Foundation)
 Mohamed Mokbel (University of Minnesota - Twin Cities)
 Mohammad Sadoghi (University of California, Davis)
 Mohammad Javad Amiri (Stony Brook University)
 Mohammed J Saeed (Apple)
 Mostafa Milani (The University of Western Ontario)
 Nick Koudas (University of Toronto)
 Nikolay Yakovets (TU Eindhoven)
 Oktie Hassanzadeh (IBM Research)
 Oliver A Kennedy (University at Buffalo, SUNY)
 Oscar Romero (Universitat Politècnica de Catalunya)
 Panagiotis Bouros (Johannes Gutenberg University Mainz)
 Panagiotis Karras (University of Copenhagen)
 Paolo Garza (Politecnico di Torino)
 Paolo Merialdo (Università degli Studi Roma Tre)
 Paolo Missier (University of Birmingham)
 Paolo Papotti (EURECOM)
 Patrick Damme (Technische Universität Berlin)
 Patrick Schäfer (Humboldt-Universität zu Berlin)
 Paul Boniol (Inria, Ecole normale supérieure)
 Peizhi Wu (University of Pennsylvania)
 Peng Peng (Hunan University)
 Periklis Chrysogelos (Oracle)
 Peter M. Fischer (University of Augsburg)
 Pierangela Samarati (Università degli Studi di Milano)
 Prashant Pandey (Northeastern University)
 Protiva Rahman (University of Florida)
 Qingqing Ye (Hong Kong Polytechnic University)
 Qitong Wang (Harvard University)
 Qizhen Zhang (University of Toronto)
 Rada Chirkova (NC State University)
 Rainer Gemulla (Universität Mannheim)

Rana Alotaibi (King Abdulaziz City for Science and Technology (KACST))
 Rana Shahout (Harvard)
 Renata Borovica-Gajic (University of Melbourne)
 Renchi Yang (Hong Kong Baptist University)
 Ria Borromeo (University of Philippines)
 Riccardo Torlone (Roma Tre University)
 Rihan Hai (TU Delft)
 Ritesh Sarkhel (Amazon)
 Roei Shraga (WPI)
 Romila Pradhan (Purdue University)
 Ruiyuan Li (Chongqing University)
 Ryan Marcus (University of Pennsylvania)
 Ryan McKenna (University of Massachusetts Amherst)
 Sai Wu (Zhejiang University)
 Sara Foresti (Universita' degli Studi di Milano)
 Sarah Kleest-Meißner (Hasselt University)
 Sebastian Link (University of Auckland)
 Sebastian Michel (RPTU Kaiserslautern-Landau)
 Sepanta Zeighami (University of California Berkeley)
 Shantanu Sharma (New Jersey Institute of Technology)
 Shaoxu Song (Tsinghua University)
 Sharad Mehrotra (U.C. Irvine)
 Shen Liang (Université Paris Cité)
 Shihabur Chowdhury (Apple)
 Shimin Di (The Hong Kong University of Science and Technology)
 Silviu Maniu (Université Grenoble Alpes)
 Siqiang Luo (Nanyang Technological University)
 Søren Keiser Jensen (Aalborg University)
 Soror Sahri (Université Paris Cité)
 Stavros Maroulis (ATHENA Research Center)
 Stefania Dumbrava (ENSIIE, INRIA, IRIF, Télécom-SudParis)
 Stefano Marchesin (Università di Padova)
 Stefano Paraboschi (Universita' degli Studi di Bergamo)
 Steffen Zeuch (TU Berlin)
 Surajit Chaudhuri (Microsoft)
 Suyash Gupta (University of Oregon)
 Sven Groppe (Universität zu Lübeck/Institute of Information Systems (IFIS))
 Tanzima Hashem (Bangladesh University of Engineering and Technology)
 Tarique Ashraf Siddiqui (Microsoft Corporation)
 Tenindra Abeywickrama (RIKEN Center for Computational Science)
 Theodoros Chondrogiannis (Norwegian University of Science and Technology (NTNU))
 Thomas Neumann (TUM)
 Tianyi Li (Aalborg University)
 Tianzheng Wang (Simon Fraser University)
 Tieying Zhang (Bytedance)
 Ting Yu (MBZUAI)
 Tingjian Ge (University of Massachusetts, Lowell)
 Tong Chen (The University of Queensland)
 Torben Bach Pedersen (Aalborg University)
 Tristan Allard (Univ Rennes, CNRS, IRISA)
 Tsz Nam Chan (Shenzhen University)
 Ulf Leser (Humboldt-Universität zu Berlin)
 Utku Sirin (Harvard University)
 Venkata Vamsikrishna Meduri (IBM Research-Almaden)

Verena Kantere (University of Ottawa)
 Vikram Nathan (Amazon, Inc.)
 Viktor Leis (Technische Universität München)
 Vivek Narasayya (Microsoft)
 Vivek Shah (Samsung)
 Wang-Chien Lee (Pennsylvania State University, USA)
 Wei Dong (Nanyang Technological University)
 WEI LU (Renmin University of China)
 Weiguo Zheng (Fudan University)
 Wenqi FAN (The Hong Kong Polytechnic University)
 Wentao Wu (Microsoft Research)
 Wentao Zhang (Peking University)
 William Schultz (MongoDB)
 Xiang Lian (Kent State University)
 Xiangyao Yu (University of Wisconsin-Madison)
 Xiangyu Ke (Zhejiang University)
 Xiao Qin (Snowflake)
 Xiao Yan (Wuhan University)
 Xiaofei Zhang (University of Memphis)
 Xiaohui Yu (York University)
 Xiaou Ding (Harbin Institute of Technology)
 Xiaoying Wang (Microsoft)
 Xike Xie (University of Science and Technology of China)
 Xin Huang (Hong Kong Baptist University)
 Xuhao Chen (MSU)
 Xun Jian (Northwestern Polytechnical University)
 Xupeng Miao (Purdue University)
 Yael Amsterdamer (Bar-Ilan University)
 Yan Huang (University of North Texas)
 Yang Cao (Institute of Science Tokyo)
 Yang Wang (The Ohio State University)
 YAO LU (NUS)
 Yesdaulet Izenov (Nazarbayev University)
 Yeye He (Microsoft Research)
 Yifan Wang (University of Hawaii)
 Yihao Ang (National University of Singapore)
 Yin Yang (Hamad bin Khalifa University)
 Yinan Li (Microsoft Research)
 Yinghui Wu (Case Western Reserve University)
 Yingtai Xiao (TikTok)
 Yiwen Zhu (Microsoft)
 Yixiang Fang (School of Data Science, The Chinese University of Hong Kong, Shenzhen)
 Yongjoo Park (UIUC)
 Yongluan Zhou (University of Copenhagen)
 Yongxin Tong (Beihang University)
 Yu Yang (City University of Hong Kong)
 Yuan Qiu (Southeast University)
 Yujie Hui (Google)
 Yuke Wang (Rice University)
 Yuyu Luo (HKUST (GZ))
 Zeyu Ding (Binghamton University)
 Zhao Cao (Remin University of China)
 ZHAOJING LUO (Beijing Institute of Technology)
 Zhengjie Miao (Simon Fraser University)
 Zhewei Wei (Renmin University of China)
 ZHIWEI FAN (Meta)
 Zhongle Xie (Zhejiang University)
 Zhuoyue Zhao (University at Buffalo)
 Zoi Kaoudi (IT University of Copenhagen)

Light Load Review Board

Abdulhakim Qahtan (Utrecht University)
Abolfazl Asudeh (University of Illinois Chicago)
Ahmed El-Roby (Carleton University)
Alberto Abelló (Universitat Politècnica de Catalunya)
Alex Conway (Cornell Tech)
Alexander Erben (Technical University of Munich)
Alexandros Labrinidis (University of Pittsburgh)
Anupam Sanghi (TU Darmstadt)
Ariful Azad (Texas A&M University)
Babak Salimi (University of California at San Diego)
Baotong Lu (Microsoft Research)
Bettina Kemme (McGill University)
C. Seshadhri (UCSC)
Curtis Dyreson (Utah State University)
Danfeng Zhang (Duke University)
Daniel Kocher (University of Salzburg)
Davide Mottin (Aarhus University)
Fabian Panse (University of Augsburg)
Felix Naumann (Hasso Plattner Institute, University of Potsdam)
Haitao Yuan (Nanyang Technological University)
Hangdong Zhao (Microsoft)
Hideyuki Kawashima (Keio University)
Hongyan Chang (NUS)
Hongzhi Yin (The University of Queensland)
Ingo Müller (Google)
Jiaheng Lu (University of Helsinki)
Jinho Lee (Seoul National University)
Juan Sequeda (data.world)
Katia Papakonstantinou (Athens University of Economics and Business)
Kexin Rong (Georgia Institute of Technology)
Konstantinos Mamouras (Rice University)
Kyriakos Mouratidis (Singapore Management University)
Lipeng Wan (Georgia State University)
Manolis Terrovitis (Athena Research Center)
Matteo Brucato (Microsoft Research)
Mayuresh Kunjir (Amazon AWS)
Michele Linardi (CYU)
Mo Sha (Alibaba Group)
Murat Demirbas (MongoDB Research)
Myeongjae Jeon (POSTECH)
Nikos Giatrakos (Technical University of Crete)
Odysseas Papapetrou (TU Eindhoven)
Panagiotis Liakos (Athens University of Economics and Business)
Panagiotis Liakos (University of Athens)
Panos Vassiliadis (University of Ioannina)
Paris Carbone (KTH Royal Institute of Technology)
Paul Groth (University of Amsterdam)
Pierre Bourhis (CNRS)
Quang-Trung Ta (National University of Singapore)
Rathijit Sen (Microsoft)
Remy Wang (University of California, Los Angeles)
Renzo Angles (Universidad de Talca)
Seokki Lee (University of Cincinnati)
Sergi Nadal (Universitat Politècnica de Catalunya)
Sivaprasad Sudhir (MIT)
Song Jiang (University of Texas, Arlington)

Sonia Bergamaschi (Università di Modena e Reggio Emilia)
Subarna Chatterjee (Datastax)
Subhadeep Sarkar (Brandeis University)
Sujaya Maiyya (University of Waterloo)
Supawit Chockchowwat (University of Illinois Urbana-Champaign)
Theodore Johnson (AT&T Labs - Research)
Thomas Fahringer (University of Innsbruck)
Thorsten Papenbrock (Philipps University of Marburg)
Vassilis J. Tsotras (UC Riverside)
Venkatesh Emani (Microsoft)
Wenjia He (University of Michigan)
Wolfgang Gatterbauer (Northeastern University)
Xingguang Chen (National University of Singapore)
Yogesh Simmhan (Indian Institute of Science (IISc), Bangalore)
Yuxin Tang (Rice University)
Zhe Jiang (University of Florida)
Ziawasch Abedjan (BIFOLD/TU Berlin)

LETTER FROM THE EDITORS IN CHIEF

We are pleased to introduce Issue 9 of Volume 19 of the Proceedings of the VLDB Endowment.

This issue brings together 50 papers spanning a broad range of topics across modern data management, including Data Mining and Analytics (8), Machine Learning, AI, and Databases (6), Information Integration and Data Quality (7), Database Engines (7), Graph and Network Data (4), Novel Database Architectures (4), Data Privacy and Security (3), Database Performance and Manageability (2), Distributed Database Systems (2), Specialized and Domain-Specific Data Management (2), Text and Semi-Structured Data (2), Provenance and Workflows (1), User Interfaces (1), and Languages (1).

A prominent theme in this issue is the growing interaction between data management and AI. Several papers study the design, optimization, and evaluation of systems involving LLMs, compound AI serving, text-to-SQL, retrieval-augmented generation, knowledge graph evaluation, document understanding, and unstructured data analysis. Together, these papers reflect the community's increasing attention to how database techniques can support, evaluate, and improve AI-driven data applications.

This issue also features strong contributions on core database systems and architectures. The papers address query optimization, transaction processing, LSM-tree-based storage, learned compaction, vector search, hashing, concurrency control on disaggregated memory, GPU query processing, and high-performance I/O. These works continue PVLDB's tradition of advancing the foundations and performance of data-intensive systems while responding to emerging hardware and deployment environments.

The issue further includes a substantial set of contributions on information integration and data quality. These papers address data valuation for machine learning, the discovery of approximate data constraints, on-demand data fusion in dynamic data lakes, and human-centered exploration of table unionability. They also extend data quality and integration challenges to knowledge graphs and unstructured data, with work on LLM-assisted knowledge graph accuracy evaluation, multimodal knowledge graph completion, and systematic benchmarking of LLM-based unstructured data analysis. Together, these contributions highlight the growing breadth of methods for assessing, integrating, enriching, and extracting value from heterogeneous data.

Finally, the issue includes timely work on privacy, security, fairness, and trustworthy data management, including secure join processing, rollback detection for trusted distributed computing, auditing demographic bias in opaque rankings, and measuring database unfairness under differential privacy. These papers underscore the importance of building data systems that are not only efficient and scalable, but also secure, reliable, and socially responsible.

We thank the authors for submitting their work to PVLDB, the reviewers and meta-reviewers for their careful and constructive evaluations, and the broader PVLDB community for its continued commitment to rigorous and impactful research. We hope that readers will find this issue both technically rich and reflective of the evolving challenges and opportunities in modern data management.

Sincerely,

Yanlei Diao and Xiaokui Xiao
Editors-in-Chief of PVLDB Vol. 19
Program Chairs for VLDB 2026