



Proceedings of the VLDB Endowment

Volume 17, No. 12– August 2024

Editors in Chief:

Meihui Zhang and Cyrus Shahabi

Associate Editors:

Alkis Polyzotis, Amol Deshpande, Angela Bonifati, Ashraf Abounnaga, Ashwin Machanavajjhala,
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. 12, August 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 12, August 2024

Pages i – xix and 3720 - 4556

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	xi
Industrial Track Chairs and Reviewers – Vol. 17	xiv
Demonstration Track Chairs and Reviewers – Vol. 17	xv
Tutorial Track Chairs and Reviewers – Vol. 17	xvi

Industrial Papers

Cloud Actor-Oriented Database Transactions in Orleans.....	3720
<i>Tamer Eldeeb, Sebastian Burckhardt, Reuben Bond, Asaf Cidon, Junfeng Yang, Philip A. Bernstein</i>	
ClickHouse - Lightning Fast Analytics for Everyone	3731
<i>Robert Schulze, Tom Schreiber, Ilya Yatsishin, Ryadh Dahimene, Alexey Milovidov</i>	
PALF: Replicated Write-ahead Logging for Distributed Databases	3745
<i>Fusheng Han, Hao Liu, Bin Chen, Debin Jia, Jianfeng Zhou, Xuwang Teng, Chuanhui Yang, Huafeng Xi, Wei Tian, Shuning Tao, Sen Wang, Quanqing Xu, Zhenkun Yang</i>	
Towards Resource Efficiency: Practical Insights into Large-Scale Spark Workloads at ByteDance	3759
<i>Yixin Wu, Xiuqi Huang, Wei Zhongjia, Hang Cheng, Chaohui Xin, Zuzhi Chen, Binbin Chen, Yufei Wu, Hao Wang, Tieying Zhang, Rui Shi, Xiaofeng Gao, Yuming Liang, Pengwei Zhao, Guihai Chen</i>	
SingleStore-V: An Integrated Vector Database System in SingleStore.....	3772
<i>Cheng Chen, Chenzhe Jin, Yunan Zhang, Sasha Podolsky, Chun Wu, Szu-Po Wang, Eric Hanson, Zhou Sun, Robert Walzer, Jianguo Wang</i>	
GaussDB: A Cloud-Native Multi-Primary Database with Compute-Memory-Storage Disaggregation	3786
<i>Guoliang Li, Wengang Tian, Jinyu Zhang, Ronen Grosman, Zongchao Liu, Sihao Li</i>	
LavaStore: ByteDance’s Purpose-built, High-performance, Cost-effective Local Storage Engine for Cloud Services	3799
<i>Hao Wang, Jiaxin Ou, Ming Zhao, Sheng Qiu, Yizheng Jiao, Yi Wang, Qizhong Mao, Zhengyu Yang, Yang Liu, Jianyang Hu, Jingwei Zhang, Jinrui Liu, Jiaqiang Chen, Yong Sheng, Cao Lixun, Heng Zhang, Hongde Li, Ming Li, Yue Ma, Lei Zhang, Jian Liu, Guanghui Zhang, Fei Liu, Jianjun Chen</i>	
Membrane - Safe and Performant Data Access Controls in Apache Spark in the Presence of Imperative Code	3813
<i>Andrei Paduroiu, Sungheun Wi, Yan Yan, Roni Burd, Ruhollah Farchtchi, Giovanni Matteo Fumarola</i>	
An Examination of CXL Memory Use Cases for In-Memory Database Management Systems using SAP HANA	3827
<i>Minseon Ahn, Thomas Willhalm, Norman May, Donghun Lee, Suprasad Mutalik Desai, Daniel Booss, Jungmin Kim, Navneet Singh, Daniel Ritter, Oliver Rebholz</i>	
KGFabric: A Scalable Knowledge Graph Warehouse for Enterprise Data Interconnection	3841
<i>Peng Yi, Lei Liang, Zhang Da, Chen Yong, Jinye Zhu, Xiangyu Liu, Kun Tang, Jialin Chen, Hao Lin, Leijie Qiu, Jun Zhou</i>	

Db2une: Tuning Under Pressure via Deep Learning	3855
<i>Alexander Bianchi, Andrew Chai, Vincent Corvinelli, Parke Godfrey, Jarek Szlichta, Calisto Zuzarte</i>	
TDSQL: Tencent Distributed Database System	3869
<i>Yuxing Chen, Anqun Pan, Hailin Lei, Anda Ye, Shuo Han, Yan Tang, Wei Lu, Yunpeng Chai, Feng Zhang, Xiaoyong Du</i>	
A Flexible Forecasting Stack	3883
<i>Tim Januschowski, Yuyang Wang, Jan Gasthaus, Syama Rangapuram, Caner Turkmen, Jasper Zschiegner, Lorenzo Stella, Michael Bohlke-Schneider, Danielle Maddix, Konstantinos Benidis, Alexander Alexandrov, Christos Faloutsos, Sebastian Schelter</i>	
Galaxybase: A High Performance Native Distributed Graph Database for HTAP	3893
<i>Bing Tong, Yan Zhou, Chen Zhang, Jianheng Tang, Jing Tang, Leihong Yang, Qiye Li, Manwu Lin, Zhongxin Bao, Jia Li, Lei Chen</i>	
SecuDB: An In-enclave Privacy-preserving and Tamper-resistant Relational Database	3906
<i>Xinying Yang, Cong Yue, Wenhui Zhang, Yang Liu, Beng Chin Ooi, Jianjun Chen</i>	
AutoTQA: Towards Autonomous Tabular Question Answering through Multi-Agent Large Language Models	3920
<i>Jun-Peng Zhu, Peng Cai, Kai Xu, Li Li, Yishen Sun, Shuai Zhou, Haihuang Su, Liu Tang, Qi Liu</i>	
ResLake: Towards Minimum Job Latency and Balanced Resource Utilization in Geo-distributed Job Scheduling.....	3934
<i>XinChun Zhang, Aqsa Kashaf, Yihan Zou, Wei Zhang, Weibo Liao, Song Haoxiang, Jintao Ye, Yakun Li, Rui Shi, Yong Tian, Feng Wei, Binbin Chen, Zuzhi Chen, Tiejing Zhang, Yongping Tang</i>	
Adaptive and Robust Query Execution for Lakehouses At Scale.....	3947
<i>Maryann Xue, Yingyi Bu, Abhishek Somani, Wenchen Fan, Ziqi Liu, Steven Chen, Herman Van Hovell, Bart Samwel, Mostafa Mokhtar, Rk Korlapati, Andy Lam, Yunxiao Ma, Vuk Ercegovic, Jiexing Li, Alexander Behm, Yuanjian Li, Xiao Li, Sriram Krishnamurthy, Amit Shukla, Michalis Petropoulos, Sameer Paranjpye, Reynold Xin, Matei Zaharia</i>	
Transparent Migration from Datastore to Firestore	3960
<i>Ed Davisson, Tilo Dickopp, David Gay, Eric Karasuda, Ram Kesavan, Vadim Yushprakh</i>	
Complex-Path: Effective and Efficient Node Ranking with Paths in Billion-Scale Heterogeneous Graphs	3973
<i>Jinquan Hang, Zhiqing Hong, Xinyue Feng, Guang Wang, Dongjiang Cao, Jiayang Qiao, Haotian Wang, Desheng Zhang</i>	
SecretFlow-SCQL: A Secure Collaborative Query pLatform	3987
<i>Wenjing Fang, Shunde Cao, Guojin Hua, Junming Ma, Yongqiang Yu, Qunshan Huang, Jun Feng, Jin Tan, Xiaopeng Zan, Pu Duan, Yang Yang, Li Wang, Ke Zhang, Lei Wang</i>	
Towards Millions of Database Transmission Services in the Cloud	4001
<i>Hua Fan, Dachao Fu, Xu Wang, Jiachi Zhang, Chaoji Zuo, Zhengyi Wu, Miao Zhang, Kang Yuan, Xizi Ni, Huo Guocheng, Wenchao Zhou, Feifei Li, Jingren Zhou</i>	
Large-Scale Metric Computation in Online Controlled Experiment Platform	4014
<i>Tao Xiong, Yong Wang</i>	
X-Stor: A Cloud-native NoSQL Database Service with Multi-model Support.....	4025
<i>Hongyu Lei, Chunhua Li, Ke Zhou, Jianping Zhu, Kezhou Yan, Fen Xiao, Ming Xie, Jiang Wang, Shiyu Di</i>	

Resource Management in Aurora Serverless	4038
<i>Bradley Barnhart, Marc Brooker, Daniil Chinenkov, Tony Hooper, Jihoun Im, Prakash Chandra Jha, Tim Kraska, Ashok Kurakula, Alexey Kuznetsov, Grant Mcalister, Arjun Muthukrishnan, Aravinthan Narayanan, Douglas Terry, Bhuvan Uргаonkar, Jiaming Yan</i>	
SQL has problems. We can fix them: Pipe syntax in SQL	4051
<i>Jeff Shute, Shannon Bales, Matthew Brown, Jean-Daniel Browne, Brandon Dolphin, Romit Kudtarkar, Andrey Litvinov, Jingchi Ma, John D Morcos, Michael Shen, David Wilhite, Xi Wu, Lulan Yu</i>	
Apache TsFile: An IoT-native Time Series File Format	4064
<i>Xin Zhao, Jialin Qiao, Xiangdong Huang, Chen Wang, Shaoxu Song, Jianmin Wang</i>	
Presto’s History-based Query Optimizer	4077
<i>Pranjal Shankhdhar, Feilong Liu, Jay Narale, James Sun, Rebecca Schlussel, Lyublena Antova</i>	
OptScaler: A Collaborative Framework for Robust Autoscaling in the Cloud	4090
<i>Ding Zou, Wei Lu, Zhibo Zhu, Xingyu Lu, Jun Zhou, Xiaojin Wang, Kangyu Liu, Kefan Wang, Renen Sun, Haiqing Wang</i>	
Dealing with Acronyms, Abbreviations, and Typos in Real-World Entity Matching.....	4104
<i>Joshua Wu, Dixin Tang, Nithin V Chalapathi, Tristan Chambers, Julie Ciccolini, Cheryl Phillips, Lisa Pickoff-White, Aditya Parameswaran</i>	
Lindorm-UWC: An Ultra-Wide-Column Database for Internet of Vehicles.....	4117
<i>Qianyu Ouyang, Chunhui Shen, Wenlong Yang, Peng Yu, Qiang Xiao, Jianhui Lei, Yadong Chen, Qilu Zhong, Xiang Wang, Yong Lin, Qingyi Meng, Zhicheng Ji, Wei Meng, Cen Zheng, Sheng Wang, Dan Pei, Wei Zhang, Feifei Li, Jingren Zhou</i>	
DLRover-RM: Resource Optimization for Deep Recommendation Models Training in the Cloud.....	4130
<i>Qinlong Wang, Tingfeng Lan, Yinghao Tang, Bo Sang, Ziling Huang, Yiheng Du, Haitao Zhang, Jian Sha, Hui Lu, Yuanchun Zhou, Ke Zhang, Mingjie Tang</i>	
Differentially Private Stream Processing at Scale	4145
<i>Bing Zhang, Vadym Doroshenko, Peter Kairouz, Thomas Steinke, Abhradeep Thakurta, Ziyin Ma, Eidan Cohen, Himani Apte, Jodi Spacek</i>	
Petabyte-Scale Row-Level Operations in Data Lakehouses.....	4159
<i>Anton Okolnychyi, Chao Sun, Kazuyuki Tanimura, Russell Spitzer, Ryan Blue, Szehon Ho, Yufei Gu, Vishwanath Lakkundi, Db Tsai</i>	
SPADE: Synthesizing Data Quality Assertions for Large Language Model Pipelines	4173
<i>Shreya Shankar, Haotian Li, Parth Asawa, Madelon Hulsebos, Yiming Lin, J.D. Zamfirscu-Pereira, Harrison Chase, William Fu-Hinthorn, Aditya Parameswaran, Eugene Wu</i>	
Simple (yet Efficient) Function Authoring for Vectorized Engines	4187
<i>Laith Sakka, Pedro Pedreira, Orri Erling, Masha Basmanova, Kevin Wilfong, Wei He, Xiaoxuan Meng, Krishna Pai, Bikramjeet Vig</i>	
Grouping, Subsumption, and Disjunctive Join Optimizations in Oracle	4200
<i>Rafi Ahmed, Krishna Kantikiran Pasupuleti, Sriram Tirupattur, Lei Sheng, Hong Su, Mohamed Ziauddin</i>	

Tutorials

LLM for Data Management	4213
<i>Guoliang Li, Xuanhe Zhou, Xinyang Zhao</i>	
Native Distributed Databases: Problems, Challenges and Opportunities	4217
<i>Quanqing Xu, Chuanhui Yang, Aoying Zhou</i>	
A Reproducible Tutorial on Reproducibility in Database Systems Research	4221
<i>Tim Fischer, Denis Hirn, Gokhan Kul</i>	
Fairness in Preference Queries: Social Choice Theories Meet Data Management.....	4225
<i>Senjuti Basu Roy, Baruch Schieber, Nimrod Talmon</i>	
Time-Series Anomaly Detection: Overview and New Trends	4229
<i>Qinghua Liu, Paul Boniol, Themis Palpanas, John Paparrizos</i>	
Consensus in Data Management With Use Cases in Edge-Cloud and Blockchain Systems	4233
<i>Faisal Nawab, Mohammad Sadoghi</i>	
Efficient Training of Graph Neural Networks on Large Graphs.....	4237
<i>Yanyan Shen, Lei Chen, Jingzhi Fang, Xin Zhang, Shihong Gao, Hongbo Yin</i>	
Workload Placement on Heterogeneous CPU-GPU Systems.....	4241
<i>Marcos N. L. Carvalho, Alkis Simitsis, Anna Queralt, Oscar Romero</i>	
Spatial Query Optimization With Learning	4245
<i>Xin Zhang, Ahmed Eldawy</i>	
Composable Data Management: An Execution Overview.....	4249
<i>Pedro Pedreira, Deepak Majeti, Orri Erling</i>	

Demonstrations

Spade: A Real-Time Fraud Detection Framework	4253
<i>Jiaxin Jiang, Zhen Zhang, Bingqiao Luo, Bingsheng He, Min Chen, Wei Yang Wang, Jia Chen</i>	
A Data-driven Spatiotemporal Simulator for Reinforcement Learning Methods	4257
<i>Dingyuan Shi, Bingchen Song, Yuanyuan Zhang, Haolong Yang, Ke Xu</i>	
BFTGym: An Interactive Playground for BFT Protocols	4261
<i>Haoyun Qin, Chenyuan Wu, Mohammad Javad Amiri, Ryan Marcus, Boon Thau Loo</i>	
DTGraph: Declarative Transformations of Property Graphs	4265
<i>Angela Bonifati, Yann Ramusat, Filip Murlak, Amela Fejza, Rachid Echahed</i>	
MLOS in Action: Bridging the Gap Between Experimentation and Auto-Tuning in the Cloud	4269
<i>Brian Kroth, Sergiy Matuskevych, Rana Alotaibi, Yiwen Zhu, Anja Gruenheid, Yuanyuan Tian</i>	
Snapcase - Regain Control over Your Predictions with Low-Latency Machine Unlearning.....	4273
<i>Sebastian Schelter, Stefan Grafberger, Maarten de Rijke</i>	

DiversiNews: Enriching News Consumption with Relevant yet Diverse News Articles Retrieval.....	4277
<i>Yiqun Sun, Qiang Huang, Yanhao Wang, Anthony K. H. Tung</i>	
SpannerLib: Embedding Declarative Information Extraction in an Imperative Workflow	4281
<i>Dean Light, Ahmad Aiashi, Mahmoud Diab, Daniel Nachmias, Stijn Vansummeren, Benny Kimelfeld</i>	
Demonstrating TabEE: Tabular Embedding Explanations	4285
<i>Roni Copul, Nave Frost, Tova Milo, Kathy Razmadze</i>	
Navigating Data Repositories: Utilizing Line Charts to Discover Relevant Datasets.....	4289
<i>Daomin Ji, Hui Luo, Zhifeng Bao, Shane Culpepper</i>	
Graph Association Analyses for Early Drug Discovery	4293
<i>Wenfei Fan, Daji Li, Peiyu Liang, Shuhao Liu, Yaoshu Wang, Yiming Wang, Min Xie, Runjie Zhang</i>	
Demonstration of MaskSearch: Efficiently Querying Image Masks for Machine Learning Workflows	4297
<i>Linsey Linxi Wei, Chung Yik Edward Yeung, Hongjian Yu, Jingchuan Zhou, Dong He, Magdalena Balazinska</i>	
Looking Deeply into the Magic Mirror: An Interactive Analysis of Database Index Selection Approaches	4301
<i>Stefan Halfpap, Jan Kossmann, Rainer Schlosser, Volker Markl</i>	
UTOPIA: Automatic Pivot Table Assistant.....	4305
<i>Whanhee Cho, Anna Fariha</i>	
TSGAssist: An Interactive Assistant Harnessing LLMs and RAG for Time Series Generation Recommendations and Benchmarking.....	4309
<i>Yihao Ang, Yifan Bao, Qiang Huang, Anthony K. H. Tung, Zhiyong Huang</i>	
DoppelGanger++ in Action: A Database Replay System with Fast Dependency Graph Generation	4313
<i>Wonseok Lee, Jaehyun Ha, Wook-Shin Han, Changgyoo Park, Myunggon Park, Juhyeng Han</i>	
LucidScript: Bottom-up Standardization for Data Preparation.....	4317
<i>Eugenie Y. Lai, Yuze Lou, Brit Youngmann, Michael Cafarella</i>	
OSSInsight: Scalable GitHub Analysis	4321
<i>Ahmad Ghazal, Zhiyuan Liang, Sunny Bains, Hanumath Maduri</i>	
IsoVista: Black-box Checking Database Isolation Guarantees.....	4325
<i>Long Gu, Si Liu, Tiancheng Xing, Hengfeng Wei, Yuxing Chen, David A Basin</i>	
ImputeVIS: An Interactive Evaluator to Benchmark Imputation Techniques for Time Series Data	4329
<i>Mourad Khayati, Quentin Nater, Jacques Pasquier</i>	
An Interactive Multi-modal Query Answering System with Retrieval-Augmented Large Language Models	4333
<i>Mengzhao Wang, Haotian Wu, Xiangyu Ke, Yunjun Gao, Xiaoliang Xu, Lu Chen</i>	
DBG-PT: A Large Language Model Assisted Query Performance Regression Debugger	4337
<i>Victor Giannakouris, Immanuel Trummer</i>	
Rodeo: Making Refinements for Diverse Top-k Queries.....	4341
<i>Felix S. Campbell, Julia Stoyanovich, Yuval Moskovitch</i>	

QPJVis Demo: Quality-boost Progressive Join Query Processing System.....	4345
<i>Xin Zhang, Ahmed Eldawy</i>	
Counterfactual Explanation Analytics: Empowering Lay Users to Take Action Against Consequential Automated Decisions	4349
<i>Peter M. Vannostrand, Dennis M. Hofmann, Lei Ma, Belisha Genin, Randy Huang, Elke A. Rundensteiner</i>	
UniView: A Unified Autonomous Materialized View Management System for Various Databases.....	4353
<i>Zhenrong Xu, Pengfei Wang, Guoze Xue, Qitong Yan, Shenghao Gong, Yelan Jiang, Yuren Mao, Yunjun Gao, Shu Shen, Wei Zhang, Dan Luo, Lu Chen</i>	
A Demonstration of TENDS: Time Series Management System based on Model Selection	4357
<i>Yuanyuan Yao, Shenjia Dai, Yilin Li, Lu Chen, Dimeng Li, Yunjun Gao, Tianyi Li</i>	
SEER: An End-to-End Toolkit for Benchmarking Time Series Database Systems in Monitoring Applications	4361
<i>Luca Althaus, Mourad Khayati, Abdelouahab Khelifati, Anton Dignös, Djellel Difallah, Philippe Cudre-Mauroux</i>	
Demonstration of DB-GPT: Next Generation Data Interaction System Empowered by Large Language Models	4365
<i>Siqiao Xue, Danrui Qi, Caigao Jiang, Fangyin Cheng, Keting Chen, Zhiping Zhang, Hongyang Zhang, Ganglin Wei, Wang Zhao, Fan Zhou, Hong Yi, Shaodong Liu, Hongjun Yang, Faqiang Chen</i>	
DeepSketch: A Query Sketching Interface for Deep Time Series Similarity Search.....	4369
<i>Zheng Zhang, Joey Shao, Andrew Crotty</i>	
Rock: Cleaning Data with both ML and Logic Rules.....	4373
<i>Zian Bao, Binbin Bie, Wenfei Fan, Daji Li, Mengyun Li, Kaiwen Lin, Wei Lin, Peijie Liu, Peng Liu, Zhicong Lv, Mingliang Ouyang, Chenyang Sun, Shuai Tang, Yaoshu Wang, Qiyuan Wei, Xiangqian Wu, Min Xie, Jing Zhang, Zhao Runxiao, Jie Zhu, Yilin Zhu</i>	
Clean4TSDB: A Data Cleaning Tool for Time Series Databases.....	4377
<i>Xiaou Ding, Song Yichen, Hongzhi Wang, Donghua Yang, Chen Wang, Jianmin Wang</i>	
LakeCompass: An End-to-End System for Table Maintenance, Search and Analysis in Data Lakes	4381
<i>Chengliang Chai, Yuhao Deng, Yutong Zhan, Ziqi Cao, Yuanfang Zhang, Lei Cao, Yu-Ping Wang, Zhiwei Zhang, Ye Yuan, Guoren Wang, Nan Tang</i>	
DOP-SQL: A General-purpose, High-utility, and Extensible Private SQL System	4385
<i>Jianzhe Yu, Wei Dong, Juanru Fang, Dajun Sun, Ke Yi</i>	
Catcher: A Cache Analysis System for Top-k Pub/Sub Service	4389
<i>Baolong Mei, Yafei Li, Wei Chen, Linshen Luan, Guanglei Zhu, Yuanyuan Jin, Jianliang Xu</i>	
Optimizing Distributed Tiered Data Storage Systems with DITIS	4393
<i>Sotiris Vasileiadis, Matthew Paraskeva, George Savva, Andreas Efstathiou, Edson Ramiro Lucas Filho, Jianqiang Shen, Lun Yang, Kebo Fu, Herodotos Herodotou</i>	
VQFT: A Visual Query Approach Based on Full-Text Search for Knowledge Graphs.....	4397
<i>Zhaozhuo Li, Xin Wang, Meng Wang, Yajun Yang, Bohan Li, Dong Han</i>	
CORAL: Collaborative Automatic Labeling System based on Large Language Models	4401
<i>Zhen Zhu, Yibo Wang, Shouqing Yang, Lin Long, Runze Wu, Xiu Tang, Junbo Zhao, Haobo Wang</i>	

CMixing: An Efficient Coin Mixing Platform to Enhance Anonymity in Cryptocurrency Transactions	4405
<i>Wangze Ni, Yiwei Zhao, Pengze Chen, Lei Chen, Peng Cheng, Chen Zhang</i>	
CyNetDiff: A Python Library for Accelerated Implementation of Network Diffusion Models	4409
<i>Eilot W. Robson, Dhemath Reddy, Abhishek Kumar Umrawal</i>	
EncChain: Enhancing Large Language Model Applications with Advanced Privacy Preservation Techniques	4413
<i>Zhe Fu, Mo Sha, Yiran Li, Huorong Li, Yubing Ma, Sheng Wang, Feifei Li</i>	
FairEM360: A Suite for Responsible Entity Matching	4417
<i>Nima Shahbazi, Mahdi Erfanian, Abolfazl Asudeh, Fatemeh Nargesian, Divesh Srivastava</i>	
RetClean: Retrieval-Based Tabular Data Cleaning Using LLMs and Data Lakes.....	4421
<i>Zan Ahmad Naeem, Mohammad Shahmeer Ahmad, Mohamed Eltabakh, Mourad Ouzzani, Nan Tang</i>	
Mach: Firefighting Time-Critical Issues in Complex Systems Using High-Frequency Telemetry.....	4425
<i>Franco Solleza, Shihang Li, William Sun, Richard Tang, Malte Schwarzkopf, Nesime Tatbul, Andrew Crotty, David Cohen, Stan Zdonik</i>	
SketchQL Demonstration: Zero-shot Video Moment Querying with Sketches	4429
<i>Renzhi Wu, Pramod Chunduri, Dristi Shah, Ashmitha Julius Aravind, Ali Payani, Xu Chu, Joy Arulraj, Kexin Rong</i>	
DataPrice: An Interactive System for Pricing Datasets in Data Marketplaces.....	4433
<i>Zhu Yiding, Hongwei Zhang, Jiayao Zhang, Jinfei Liu, Kui Ren</i>	
Demonstration of the VeriEQL Equivalence Checker for Complex SQL Queries.....	4437
<i>Pinhan Zhao, Yang He, Xinyu Wang, Yuepeng Wang</i>	
FedSQ: A Secure System for Federated Vector Similarity Queries	4441
<i>Zeqi Zhu, Zeheng Fan, Yuxiang Zeng, Yexuan Shi, Yi Xu, Mengmeng Zhou, Jin Dong</i>	
FedSM: A Practical Federated Shared Mobility System	4445
<i>Shuyue Wei, Yuanyuan Zhang, Zimu Zhou, Tianlong Zhang, Ke Xu</i>	
DataLoom: Simplifying Data Loading with LLMs	4449
<i>Alexander Van Renen, Mihail Stoian, Andreas Kipf</i>	
Demonstration of VCR: A Tabular Data Slicing Approach to Understanding Object Detection Model Performance.....	4453
<i>Jie Jeff Xu, Saahir Dhanani, Jorge Piazzentin Ono, Wenbin He, Liu Ren, Kexin Rong</i>	
ModsNet: Performance-aware Top-k Model Search using Exemplar Datasets	4457
<i>Mengying Wang, Hanchao Ma, Sheng Guan, Yiyang Bian, Haolai Che, Abhishek Daundkar, Alp Sehirlioglu, Yinghui Wu</i>	
OFL-W3: A One-shot Federated Learning System on Web 3.0.....	4461
<i>Linshan Jiang, Moming Duan, Bingsheng He, Yulin Sun, Peishen Yan, Yang Hua, Tao Song</i>	
Swift: A Data-Driven Flight Planning System at Scale	4465
<i>Chang Gao, Tianlong Zhang, Yuxiang Zeng, Yi Xu, Shuyuan Li, Yuanyuan Zhang</i>	

Pyneapple-G: Scalable Spatial Grouping Queries	4469
<i>Laila Abdelhafeez, Andres Calderon, Amr Magdy, Vassilis J. Tsotras</i>	
PD-Explain: A Unified Python-native Framework for Query Explanations Over DataFrames	4473
<i>Itay Elyashiv, Amir Gilad, Edna Isakov, Tal Tikochinsky, Amit Somech</i>	
HocoPG: A Database System with Homomorphic Compression for Text Processing	4477
<i>Jiawei Guan, Feng Zhang, Yuxin Tang, Weitang Ye, Xiaoyong Du</i>	
Chat2Data: An Interactive Data Analysis System with RAG, Vector Databases and LLMs	4481
<i>Xinyang Zhao, Xuanhe Zhou, Guoliang Li</i>	
PrismX: A Single-Machine System for Querying Big Graphs	4485
<i>Shuhao Liu, Yang Liu, Wenfei Fan</i>	
TimeCSL: Unsupervised Contrastive Learning of General Shapelets for Explorable Time Series Analysis	4489
<i>Zhiyu Liang, Chen Liang, Zheng Liang, Hongzhi Wang, Bo Zheng</i>	
HSAP: A Human-in-the-loop Social Media-based Situation Awareness Platform	4493
<i>Xiangmin Zhou, Chengkun He, Xi Chen, Yanchun Zhang</i>	
DB-MAGS: Multi-Anomaly Data Generation System for Transactional Databases.....	4497
<i>Yiqi Shen, Sijia Li, Miaodong Shen, Peng Cai, Weiyuan Xu, Li Kai, Jinlong Cai</i>	
QuoteInspector: Gaining Insight about Social Media Discussions	4501
<i>Peizhi Wu, Yi Zhang, Wang-Chiew Tan, Zack G. Ives</i>	

Panel

Vector Databases: What's Really New and What's Next?	4505
<i>Jianguo Wang, Eric Hanson, Guoliang Li, Yannis Papakonstantinou, Harsha Simhadri, Charles Xie</i>	

Endowment Awards

High-Performance Spatial Data Analytics: Systematic R&D for Scale-Out and Scale-Up Solutions from the Past to Now	4507
<i>Fusheng Wang, Rubao Lee, Dejun Teng, Xiaodong Zhang, Joel Saltz</i>	
Intelligent Agents for Data Exploration.....	4521
<i>Sihem Amer-Yahia</i>	
Reimagining Deep Learning Systems Through the Lens of Data Systems	4531
<i>Arun Kumar</i>	
LeanStore: A High-Performance Storage Engine for NVMe SSDs	4536
<i>Viktor Leis</i>	

Keynotes

Databases Unbound: Querying All of the World's Bytes with AI	4546
<i>Samuel Madden, Michael Cafarella, Michael Franklin, Tim Kraska</i>	

Sharing Information with Differential Privacy: A Database Perspective.....	4555
<i>Xiaokui Xiao</i>	
Harmonizing ML and Databases: A Symphony of Data	4556
<i>Fatma Ozcan</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)

INDUSTRIAL TRACK CHAIRS AND REVIEWERS - Vol. 17

Industrial Track Program Chairs

Divesh Srivastava (AT&T, USA)
Nesime Tatbul (Intel Labs and MIT, USA)

Industrial Track Program Committee Members

Michael Abebe (Salesforce, Canada)
Hossein Ahmadi (Snowflake, USA)
Amirhossein Aleyasen (Datometry, USA)
Lisa Amini (IBM Research AI, USA)
Philip Bernstein (Microsoft Research, USA)
Carsten Binnig (TU Darmstadt, Germany)
Alexander Boehm (SAP SE, Germany)
Jesus Camacho-Rodriguez (Microsoft, USA)
Michael Carey (Couchbase, USA)
Jianjun Chen (Bytedance, USA)
Shimin Chen (Chinese Academy of Sciences, China)
David Cohen (Intel, USA)
Bin Cui (Peking University, China)
Sudipto Das (Amazon Web Services, USA)
A. Jesse Jiryu Davis (MongoDB, USA)
Yanlei Diao (Ecole Polytechnique and Amazon Web Services, France)
Aaron Elmore (University of Chicago, USA)
Florian Funke (Snowflake, Germany)
Georgios Giannikis (Observe Inc., Switzerland)
Wook-Shin Han (POSTECH, South Korea)
Ihab Ilyas (University of Waterloo and Apple, USA)
Alekh Jindal (SmartApps, USA)
Yaron Kanza (AT&T, USA)
Konstantinos Karanasos (Meta, USA)
Flip Korn (Google, USA)
Georgia Koutrika (Athena Research Center, Greece)
Per-Ake Larson (Huawei Research, Canada)
Wolfgang Lehner (TU Dresden, Germany)
Volker Markl (TU Berlin, Germany)
Parimarjan Negi (Databricks, USA)
Krishna Kantikiran Pasupuleti (Oracle, USA)
Jignesh Patel (Carnegie Mellon University and DataChat, USA)
Jun Rao (Confluent, USA)
Rajeev Rastogi (Amazon, India)
Mehul Shah (Aryn, USA)
Vladislav Shkapenyuk (AT&T, USA)
Yannis Sismanis (Databricks, USA)
Srini Srinivasan (Aerospike, USA)
Kian-Lee Tan (National University of Singapore, Singapore)
Wang-Chiew Tan (Meta, USA)
Nan Tang (HKUST (GZ), China)
Kapil Vaidya (Amazon Web Services, USA)
Haixun Wang (Instacart, USA)
Steven Whang (KAIST, South Korea)
Yingjun Wu (RisingWave Labs, USA)
Quanqing Xu (OceanBase, Ant Group, China)
Cong Yu (Celonis, USA)
Jingren Zhou (Alibaba Group, China)

DEMONSTRATION TRACK CHAIRS AND REVIEWERS - Vol. 17

Demonstration Track Program Chairs

Sihem Amer-Yahia (CNRS & Univ. Grenoble Alpes)
Yongxin Tong (Beihang University)
Yuncheng Wu (National University of Singapore)

Demonstration Track Program Committee Members

Alexander van Renen (UTN)
Amit Somech (Bar-Ilan University)
Andrew Crotty (Carnegie Mellon University)
Cheng Long (Nanyang Technological University)
El Kindi Rezig (University of Utah)
Enzo Veltri (Università della Basilicata)
Georgia Troullinou (FORTH-ICS)
Haoqiong Bian (Renmin University of China)
Hiroaki Shiokawa (University of Tsukuba)
Idir Benouaret (Epita Research Laboratory)
Jarek Szlichta (York University and IBM CAS)
Jianliang Xu (Hong Kong Baptist University)
Jianqiu Xu (Nanjing University of Aeronautics and Astronautics)
Jyoti Leeka (Microsoft)
Lihong He (IBM)
Makoto Onizuka (Osaka University)
Manisha Luthra (TU Darmstadt and DFKI)
Mayuresh Kunjir (Amazon AWS)
Muhammad El-Hindi (TU Darmstadt)
Nan Tang (HKUST (GZ))
Nicoleta Preda (University of Versailles)
Quanqing Xu (OceanBase, Ant Group)
Ramon Lawrence (University of British Columbia)
Roe Shraga (Northeastern University)
Shaoxu Song (Tsinghua University)
Silviu Maniu (Université Grenoble Alpes)
Tiago de Melo (Universidade do Estado do Amazonas)
Tilmann Rabl (HPI, University of Potsdam)
Tsz Nam Chan (Shenzhen University)
Varun Pandey (BIFOLD, TU Berlin)
Venkatesh Emani (Microsoft)
Verena Kantere (University of Ottawa)
Wangze Ni (Hong Kong University of Science and Technology)
Wenjie Zhang (University of New South Wales)
Wentao Zhang (Peking University)
Xiaochun Yang (Northeastern University)
Xin Wang (Tianjin University)
Xinying Yang (ByteDance)
Xupeng Miao (Purdue University)
Y.C. Tay (National University of Singapore)
Yaron Kanza (AT&T Labs-Research)
Yingxia Shao (BUPT)
Yiru Chen (Columbia University)
Yiwen Zhu (Microsoft)
Yu Liu (Beijing Jiaotong University)
Yunjun Gao (Zhejiang University)
Yuval Moskovitch (Ben Gurion University)
Zhanhao Zhao (National University of Singapore)

TUTORIAL TRACK CHAIRS AND REVIEWERS - Vol. 17

Tutorial Track Program Chairs

Li Xiong (Emory University, USA)

Torsten Grust (University of Tübingen, Germany)

Tutorial Track Program Committee Members

Alberto Lerner (University of Fribourg, Switzerland)

Alkis Simitsis (Athena Research Center, Greece)

Amir Shaikhha (University of Edinburgh, UK)

Chao Zhang (Tsinghua University, China)

Chenhao Ma (The Chinese University of Hong Kong, Shenzhen, China)

Chuan Xiao (Osaka University, Nagoya University, Japan)

Eamonn J. Keogh (UC Riverside, USA)

Jana Giceva (TU Munich, Germany)

Jelle Hellings (McMaster University, Canada)

Jian Pei (Duke University, USA)

Madhulika Mohanty (Inria Saclay, France)

Maurice van Keulen (University of Twente, The Netherlands)

Michael Grossniklaus (University of Konstanz, Germany)

Ramon Antonio Rodrigues Zalipynis (HSE University, Russia)

Suyash Gupta (University of California, Berkeley, USA)

Wei Wang (Hong Kong University of Science and Technology (Guangzhou), China)

Xiaolan Wang (Meta, USA)

Zoi Kaoudi (IT University of Copenhagen, Denmark)

LETTER FROM THE EDITORS IN CHIEF

We are pleased to present the twelfth issue of PVLDB (Proceedings of the VLDB) Volume 17. While the first eleven issues covered research papers accepted to PVLDB and presented at the VLDB 2024 Conference (celebrating the 50th anniversary of the VLDB conference), Issue 12 includes keynote talks, panels, workshops, as well as peer-reviewed industrial papers, demonstrations, and tutorials, all of which were also part of the VLDB 2024 program.

The Industrial Track of VLDB 2024 covered all aspects of innovative commercial or industrial-strength data management systems and deployed solutions, as well as novel real-world applications of data management systems and experience in applying recent research advances to real-world or industry-relevant problems. This year, we received 88 submissions, out of which we accepted 21 papers (24%) directly and an additional 16 papers (18%) after one round of revision. All papers were reviewed by the Industrial Track Program Committee, which consisted of 48 members with a wide range of technical expertise from all around the world. Each paper received 3 reviews, which were shared with the authors during an Author Feedback phase. Based on the reviews, the author feedback, and extensive discussion, the Industrial Track Program Chairs identified papers to be accepted, either directly or after a round of revision with guidance from a designated shepherd. The revised submissions were significant improvements over the original submissions and were all accepted based on a thorough review by one of the original reviewers acting as a shepherd for the paper. Industrial papers at this year's VLDB cover a wide range of real-world data management solutions including database engines, database technologies and management, data mining, distributed database systems, machine learning and AI for databases, and privacy and security.

The Demonstration Track of VLDB 2024 is an important platform for sharing and showcasing the latest advancements in the field of data management. The Demonstration Track received a healthy number of 166 submissions, which increased by 23% than last year. After a rigorous review process, 64 (38.5%) of these submissions were accepted. Each demonstration submission was reviewed by three reviewers. The Demonstration Program Committee responsible for the selection process was carefully assembled, taking into consideration gender and geographic diversity. This diverse committee ensured a comprehensive and inclusive evaluation of the demonstration proposals, contributing to the quality and variety of the accepted demonstrations.

The Tutorial Track of the 50th edition of VLDB covers state-of-the-art research, development, and applications in data management or related areas, including interdisciplinary areas. This year, we received 30 submissions, out of which we accepted 10 tutorials, representing an acceptance rate of 1/3. We would like to thank all authors for their contributions to the track. The Tutorial Program Committee, comprising 18 members of our community, was assembled with gender and geographic diversity in mind. Each proposal was reviewed by at least two reviewers, and then by the Tutorial Program Chairs for a final decision, with coordination with the PC Chairs and general chairs on how the proposals could complement the overall VLDB 2024 program. We are proud to say that the accepted proposals cover a wide range of topics including distributed databases and blockchain-based systems, workload and composable data management, machine learning and LLMs for data management, fairness of queries, time series data and graph data analytics, and reproducibility for database research.

VLDB 2024 received 14 workshop proposals. After reviewing the proposals and evaluating their merits based on their history, relevance of topics, and distinctiveness, we selected 11 high-quality workshops. Like previous years, we erred on the side of accepting a larger number of workshops to ensure a good diversity of topics and allow new and revived workshops to establish themselves with sufficiently many submissions.

Monday, 26th August

- 3rd International Workshop on Large Scale Graph (LSGDA)
- 13th International Workshop on Quality in Databases (QDB)
- International Workshop on Data Management Opportunities in Unifying Large Language Models + Knowledge Graphs (LLM+KG)
- 15th International Workshop on Accelerating Analytics and Data Management Systems using Modern Processor and Storage Architectures (ADMS)

Friday, 30th August

- 6th International Workshop on Foundations and Applications of Blockchain (FAB)
- 2nd Workshop on Cloud Databases (CloudDB)
- 2nd International Workshop on Tabular Data Analysis (TaDA)

- TPC Technology Conference (TPCTC)
- 2nd International Workshop on Quantum Data Science and Management (QDSM)
- 7th International Workshop on Big Data Visual Exploration and Analytics (BigVis)
- 1st International Workshop on Data-driven AI (DATAI)

The PhD Workshop features 18 student presentations, a keynote by Dr. Divesh Srivastava (AT&T) and a panel on “Success in Ph.D. Journey”.

VLDB 2024 features three keynotes:

- Databases Unbound: Querying All of the World’s Bytes with AI; *Prof. Samuel Madden*, MIT College of Computing Distinguished Professor of Computing, Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts Institute of Technology (MIT).
- Harmonizing ML and Databases: A Symphony of Data; *Dr. Fatma Ozcan*, Principal Engineer at Systems Research@Google.
- Sharing Information with Differential Privacy: A Database Perspective; *Prof. Xiaokui Xiao*, School of Computing, National University of Singapore (NUS).

VLDB 2024 also features one panel on “Vector Databases: What’s Really New and What’s Next?” moderated by Jianguo Wang (Purdue University), with panelists Eric Hanson (SingleStore), Guoliang Li (Tsinghua University), Yannis Papakonstantinou (Google Cloud), Harsha Simhadri (Microsoft), and Charles Xie (Zilliz).

Besides the Best Paper and Demonstration awards announced at the VLDB 2024 Conference, four VLDB Endowment Awards are made:

- 2024 VLDB Test of Time Award: Ablimit Aji (Emory University), Fusheng Wang (Emory University), Hoang Vo (Emory University), Rubao Lee (The Ohio State University), Qiaoling Liu (Emory University), Xiaodong Zhang (The Ohio State University) and Joel Saltz (Emory University) for their VLDB 2013 paper “Hadoop-GIS: A High Performance Spatial Data Warehousing System over MapReduce”
- 2024 VLDB Women in Database Research Award: Sihem Amer-Yahia (CNRS, Univ. Grenoble Alpes)
- 2024 VLDB Early Career Research Contribution Award: Arun Kumar (University of California, San Diego)
- 2024 VLDB Early Career Research Contribution Award: Viktor Leis (Technische Universität München)

In closing, we wish to collectively express our deep gratitude to all members of various program committees as well as our Proceedings Chairs who worked tirelessly in the past year to ensure the timely and smooth publication of PVLDB Volume 17.

Meihui Zhang, Beijing Institute of Technology
 Cyrus Shahabi, University of Southern California
Editors-in-Chief of PVLDB Vol. 17
Program Chairs for VLDB 2024

Divesh Srivastava, AT&T
 Nesime Tatbul, Intel & MIT
Industrial, Application and Experience Program Chairs

Ju Fan, Renmin University of China
 Yang Cao, Tokyo Institute of Technology
 Xiaou Ding, Harbin Institute of Technology
Proceedings Chairs

Sihem Amer-Yahia, CNRS & Univ. Grenoble Alpes
 Yongxin Tong, BUAA
 Yuncheng Wu, National University of Singapore
Demonstration Program Chairs

Li Xiong, Emory University
Torsten Grust, Universität Tübingen
Tutorial Program Chairs

Themis Palpanas, University Paris Cite
Philippe Bonnet, IT University of Copenhagen
Workshop Chairs

Haixun Wang, Instacart
Workshop Proceeding Chairs

Wook-Shin Han, Pohang University of Science and Technology
Ibrahim Sabek, University of Southern California
PhD Workshop Chairs

M. Tamer Özsu, University of Waterloo
Xiaofang Zhou, HKUST
Panel Chairs