



# Proceedings of the VLDB Endowment

Volume 16, No. 12 – August 2023

**Editors in Chief:**

Georgia Koutrika and Jun Yang

**Associate Editors:**

Alkis Simitsis, Amol Deshpande, Angela Bonifati, Ashwin Machanavajjhala, Badrish Chandramouli,  
Boris Glavic, Ce Zhang, Cyrus Shahabi, Dan Olteanu, Eric Lo, Evaggelia Pitoura, Evimaria Terzi,  
Gustavo Alonso, Helen (Zi) Huang, Hong Cheng, Kenneth Ross, Khuzaima Daudjee, Kyuseok Shim,  
Letizia Tanca, Lucian Popa, Magdalena Balazinska, Meihui Zhang, Neoklis Polyzotis,  
Nesime Tatbul, Nikos Mamoulis, Rachel Pottinger, Wenjie Zhang, Wolfgang Gatterbauer,  
Wook-Shin Han, Xiaokui Xiao, Yannis Velegrakis, Yanyan Shen, Yi Chen, Yongxin Tong, Zhifeng Bao

**Publication Editors:**

Manos Athanassoulis, Kostas Stefanidis, Ju Fan

PVLDB – Proceedings of the VLDB Endowment

Volume 16, No. 12, August 2023.

All papers published in this issue will be presented at the 49th International Conference on Very Large Data Bases, Vancouver, Canada, 2023.

## Copyright 2023 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](mailto:info@vldb.org).

Volume 16, Number 12, August 2023

Pages i – xvi and 3475 - 4159

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. 16 .....	ix
Industrial Track Chairs and Reviewers – Vol. 16 .....	xii
Demonstration Track Chairs and Reviewers – Vol. 16 .....	xiii
Tutorial Track Chairs and Reviewers – Vol. 16 .....	xiv

### Industrial Papers

Progressive Partitioning for Parallelized Query Execution in Google's Napa .....	3475
<i>Junichi Tatemura, Tao Zou, Jagan Sankaranarayanan, Yanlai Huang, Jim Chen, Yupu Zhang, Kevin Lai, Hao Zhang, Gokul Nath Babu Manoharan, Goetz Graefe, Divyakant Agrawal, Brad Adelberg, Shilpa Kolhar, Indrajit Roy</i>	
Taurus MM: bringing multi-master to the cloud .....	3488
<i>Alex Depoutovitch, Chong Chen, Per-Åke Larson, Jack Ng, Shu Lin, Guanzhu Xiong, Paul Lee, Emad Boctor, Samiao Ren, Lengdong Wu, Yuchen Zhang, Calvin Sun</i>	
StreamOps: Cloud-Native Runtime Management for Streaming Services in ByteDance.....	3501
<i>Yancan Mao, Zhanghao Chen, Yifan Zhang, Meng Wang, Yong Fang, Guanghui Zhang, Rui Shi, Richard T.B. Ma</i>	
AutoSteer: Learned Query Optimization for Any SQL Database.....	3515
<i>Christoph Anneser, Nesime Tatbul, David Cohen, Zhenggang Xu, Prithviraj Pandian, Nikolay Laptev, Ryan Marcus</i>	
Krypton: Real-time Serving and Analytical SQL Engine at ByteDance.....	3528
<i>Jianjun Chen, Rui Shi, Heng Chen, Li Zhang, Ruidong Li, Wei Ding, Liya Fan, Hao Wang, Mu Xiong, Yuxiang Chen, Benchao Dong, Kuankuan Guo, Yuanjin Lin, Xiao Liu, Haiyang Shi, Peipei Wang, Zikang Wang, Yemeng Yang, Junda Zhao, Dongyan Zhou, Zhikai Zuo, Yuming Liang</i>	
EmbedX: A Versatile, Efficient and Scalable Platform to Embed Both Graphs and High-Dimensional Sparse Data.....	3543
<i>Yuanhang Zou, Zhihao Ding, Jieming Shi, Shuting Guo, Chunchen Su, Yafei Zhang</i>	
The Story of AWS Glue.....	3557
<i>Mohit Saxena, Benjamin Sowell, Daiyan Alamgir, Nitin Bahadur, Bijay Bisht, Santosh Chandrachood, Chitti Keswani, G2 Krishnamoorthy, Austin Lee, Bohou Li, Zach Mitchell, Vaibhav Porwal, Maheedhar Reddy Chappidi, Brian Ross, Noritaka Sekiyama, Omer Zaki, Linchi Zhang, Mehul A. Shah</i>	
Towards General and Efficient Online Tuning for Spark.....	3570
<i>Yang Li, Huaijun Jiang, Yu Shen, Yide Fang, Xiaofeng Yang, Danqing Huang, Xinyi Zhang, Wentao Zhang, Ce Zhang, Peng Chen, Bin Cui</i>	
CDSBen: Benchmarking the Performance of Storage Services in Cloud-native Database System at ByteDance .....	3584
<i>Jiashu Zhang, Wen Jiang, Bo Tang, Haoxiang Ma, Lixun Cao, Zhongbin Jiang, Yuanyuan Nie, Fan Wang, Lei Zhang, Yuming Liang</i>	
FEBench: A Benchmark for Real-Time Relational Data Feature Extraction.....	3597
<i>Xuanhe Zhou, Cheng Chen, Kunyi Li, Bingsheng He, Mian Lu, Qiaosheng Liu, Wei Huang, Guoliang Li, Zhao Zheng, Yuqiang Chen</i>	

MINT: Detecting Fraudulent Behaviors from Time-series Relational Data.....	3610
<i>Fei Xiao, Yuncheng Wu, Meihui Zhang, Gang Chen, Beng Chin Ooi</i>	
Microsoft Purview: A System for Central Governance of Data .....	3624
<i>Shafi Ahmad, Dillidorai Arumugam, Srdan Bozovic, Elnata Degefa, Sailesh Duvvuri, Steven Gott, Nitish Gupta, Joachim Hammer, Nivedita Kaluskar, Raghav Kaushik, Rakesh Khanduja, Prasad Mujumdar, Gaurav Malhotra, Pankaj Naik, Nikolas Ogg, Krishna Kumar Parthasarthy, Raghu Ramakrishnan, Vlad Rodriguez, Rahul Sharma, Jakub Szymaszek, Andreas Wolter</i>	
Anser: Adaptive Information Sharing Framework of AnalyticDB .....	3636
<i>Liang Lin, Yuhan Li, Bin Wu, Huijun Mai, Renjie Lou, Jian Tan, Feifei Li</i>	
TPCx-AI - An Industry Standard Benchmark for Artificial Intelligence and Machine Learning Systems.....	3649
<i>Christoph Brücke, Philipp Härtling, Rodrigo D Escobar Palacios, Hamesh Patel, Tilmann Rabl</i>	
OneProvenance: Efficient Extraction of Dynamic Coarse-Grained Provenance From Database Query Event Logs.....	3662
<i>Fotis Psallidas, Ashvin Agrawal, Chandru Sugunan, Khaled Ibrahim, Konstantinos Karanasos, Jesús Camacho-Rodríguez, Avrielia Floratou, Carlo Curino, Raghu Ramakrishnan</i>	
Techniques and Efficiencies from Building a Real-Time DBMS .....	3676
<i>V. Srinivasan, Andrew Gooding, Sunil Sayyaparaju, Thomas Lopatic, Kevin Porter, Ashish Shinde, B. Narendran</i>	
Real-time Workload Pattern Analysis for Large-scale Cloud Databases .....	3689
<i>Jiaqi Wang, Tianyi Li, Anni Wang, Xiaozhe Liu, Lu Chen, Jie Chen, Jianye Liu, Junyang Wu, Feifei Li, Yunjun Gao</i>	
Big Data Analytic Toolkit: A general-purpose, modular, and heterogeneous acceleration toolkit for data analytical engines .....	3702
<i>Jiang Li, Qi Xie, Yan Ma, Jian Ma, Kunshang Ji, Yizhong Zhang, Chaojun Zhang, Yixiu Chen, Gangsheng Wu, Jie Zhang, Kaidi Yang, Xinyi He, Qiuyang Shen, Yanting Tao, Haiwei Zhao, Penghui Jiao, Chengfei Zhu, David Qian, Cheng Xu</i>	
Lindorm TSDB: A Cloud-native Time-series Database for Large-scale Monitoring Systems.....	3715
<i>Chunhui Shen, Qianyu Ouyang, Feibo Li, Zhipeng Liu, Longcheng Zhu, Yujie Zou, Qing Su, Tianhuan Yu, Yi Yi, Jianhong Hu, Cen Zheng, Bo Wen, Hanbang Zheng, Lunfan Xu, Sicheng Pan, Bin Wu, Xiao He, Ye Li, Jian Tan, Sheng Wang, Dan Pei, Wei Zhang, Feifei Li</i>	
OceanBase Paetica: A Hybrid Shared-nothing/Shared-everything Database for Supporting Single Machine and Distributed Cluster .....	3728
<i>Zhifeng Yang, Quanqing Xu, Shanyan Gao, Chuanhui Yang, Guoping Wang, Yuzhong Zhao, Fanyu Kong, Hao Liu, Wanhong Wang, Jinliang Xiao</i>	
SimpleTS: An Efficient and Universal Model Selection Framework for Time Series Forecasting .....	3741
<i>Yuanyuan Yao, Dimeng Li, Hailiang Jie, Lu Chen, Tianyi Li, Jie Chen, Jiaqi Wang, Feifei Li, Yunjun Gao</i>	
PolarDB-SCC: A Cloud-Native Database Ensuring Low Latency for Strongly Consistent Reads .....	3754
<i>Xinjun Yang, Yingqiang Zhang, Hao Chen, Chuan Sun, Feifei Li, Wenchao Zhou</i>	
ScalarDB: Universal Transaction Manager for Polystores.....	3768
<i>HiroYuki Yamada, Toshihiro Suzuki, Yuji Ito, Jun Nemoto</i>	
Angel-PTM: A Scalable and Economical Large-scale Pre-training System in Tencent .....	3781
<i>Xiaonan Nie, Yi Liu, Fangcheng Fu, Jinbao Xue, Dian Jiao, Xupeng Miao, Yangyu Tao, Bin Cui</i>	
Eigen: End-to-end Resource Optimization for Large-Scale Databases on the Cloud .....	3795

*Ji You Li, Jiachi Zhang, Wenchao Zhou, Yuhang Liu, Shuai Zhang, Zhuoming Xue, Ding Xu, Hua Fan, Fangyuan Zhou, Feifei Li*

MagicScaler: Uncertainty-aware, Predictive Autoscaling..... 3808  
*Zhicheng Pan, Yihang Wang, Yingying Zhang, Sean Bin Yang, Yunyao Cheng, Peng Chen, Chenjuan Guo, Qingsong Wen, Xiduo Tian, Yunliang Dou, Zhiqiang Zhou, Chengcheng Yang, Aoying Zhou, Bin Yang*

Kora: A Cloud-Native Event Streaming Platform for Kafka..... 3822  
*Anna Povzner, Prince Mahajan, Jason Gustafson, Jun Rao, Ismael Juma, Feng Min, Shriram Sridharan, Nikhil Bhatia, Gopi Attaluri, Adithya Chandra, Stanislav Kozlovski, Rajini Sivaram, Lucas Bradstreet, Bob Barrett, Dhruvil Shah, David Jacot, David Arthur, Manveer Chawla, Ron Dagostino, Colin McCabe, Manikumar Reddy Obili, Kowshik Prakasam, Jose Garcia Sancio, Vikas Singh, Alok Nikhil, Kamal Gupta*

Automatic SQL Error Mitigation in Oracle ..... 3835  
*Krishna Kantikiran Pasupuleti, Jiakun Li, Hong Su, Mohamed Ziauddin*

PyTorch FSDP: Experiences on Scaling Fully Sharded Data Parallel..... 3848  
*Yanli Zhao, Andrew Gu, Rohan Varma, Liang Luo, Chien-Chin Huang, Min Xu, Less Wright, Hamid Shojanazeri, Myle Ott, Sam Shleifer, Alban Desmaison, Can Balioglu, Pritam Damania, Bernard Nguyen, Geeta Chauhan, Yuchen Hao, Ajit Mathews, Shen Li*

## **Tutorials**

Time Series Data Mining: A Unifying View..... 3861  
*Eamonn Keogh*

Machine Learning for Subgraph Extraction: Methods, Applications and Challenges ..... 3864  
*Kai Siong Yow, Ningyi Liao, Siqiang Luo, Reynold Cheng*

Private Information Retrieval in Large Scale Public Data Repositories..... 3868  
*Ishtiyaque Ahmad, Divyakant Agrawal, Amr El Abbadi, Trinabh Gupta*

Data and AI Model Markets: Opportunities for Data and Model Sharing, Discovery, and Integration..... 3872  
*Jian Pei, Raul Castro Fernandez, Xiaohui Yu*

Efficient Execution of User-Defined Functions in SQL Queries..... 3874  
*Yannis Foufoulas, Alkis Simitsis*

Natural Language Interfaces for Databases with Deep Learning..... 3878  
*George Katsogiannis-Meimarakis, Mike Xydas, Georgia Koutrika*

Join Order Selection with Deep Reinforcement Learning: Fundamentals, Techniques, and Challenges ..... 3882  
*Zhengdong Yan, Valter Uotila, Jiaheng Lu*

Full-Power Graph Querying: State of the Art and Challenges ..... 3886  
*Ioana Manolescu, Madhulika Mohanty*

A Tutorial on Visual Representations of Relational Queries..... 3890  
*Wolfgang Gatterbauer*

Databases on Modern Networks: A Decade of Research that now comes into Practice ..... 3894  
*Alberto Lerner, Carsten Binnig, Philippe Cudré-Mauroux, Rana Hussein, Matthias Jasný, Theo Jepsen, Dan R. K. Ports, Lasse Thostrup, Tobias Ziegler*

Building a Collaborative Data Analytics System: Opportunities and Challenges..... 3898  
*Zuozhi Wang, Chen Li*

## Demonstrations

PAINE Demo: Optimizing Video Selection Queries With Commonsense Knowledge .....	3902
<i>Wenjia He, Ibrahim Sabek, Yuze Lou, Michael Cafarella</i>	
DoveDB: A Declarative and Low-Latency Video Database.....	3906
<i>Ziyang Xiao, Dongxiang Zhang, Zepeng Li, Sai Wu, Kian-Lee Tan, Gang Chen</i>	
DeepVQL: Deep Video Queries on PostgreSQL .....	3910
<i>Dong June Lew, Kihyun Yoo, Kwang Woo Nam</i>	
A Demonstration of DLBD: Database Logic Bug Detection System .....	3914
<i>Xiu Tang, Sai Wu, Dongxiang Zhang, Ziyue Wang, Gongsheng Yuan, Gang Chen</i>	
PSFQ: A Blockchain-based Privacy-preserving and Verifiable Student Feedback Questionnaire Platform .....	3918
<i>Wangze Ni, Pengze Chen, Lei Chen</i>	
QO-Insight: Inspecting Steered Query Optimizers .....	3922
<i>Christoph Anneser, Mario Petruccelli, Nesime Tatbul, David Cohen, Zhenggang Xu, Prithviraj Pandian, Nikolay Laptev, Ryan Marcus, Alfons Kemper</i>	
Lynx: A Graph Query Framework for Multiple Heterogeneous Data Sources.....	3926
<i>Zhihong Shen, Chuan Hu, Zihao Zhao</i>	
Showcasing Data Management Challenges for Future IoT Applications with NebulaStream.....	3930
<i>Aljoscha Lepping, Hoang Mi Pham, Laura Mons, Balint Rueb, Philipp M. Grulich, Ankit Chaudhary, Steffen Zeuch, Volker Markl</i>	
CEDA: Learned Cardinality Estimation with Domain Adaptation.....	3934
<i>Zilong Wang, Qixiong Zeng, Ning Wang, Haowen Lu, Yue Zhang</i>	
FastMosaic in Action: A New Mosaic Operator for Array DBMSs.....	3938
<i>Ramon Antonio Rodrigues Zalipynis</i>	
Sniffer: A Novel Model Type Detection System against Machine-Learning-as-a-Service Platforms .....	3942
<i>Zhuo Ma, Yilong Yang, Bin Xiao, Yang Liu, Xinjing Liu, Zhuoran Ma, Tong Yang</i>	
KGNav: A Knowledge Graph Navigational Visual Query System.....	3946
<i>Xiang Wang, Xin Wang, Zhaozhuo Li, Dong Han</i>	
On-the-fly Data Transformation in Action .....	3950
<i>Ju Hyoung Mun, Konstantinos Karatsenidis, Tarikul Islam Papon, Shahin Roozkhosh, Denis Hoornaert, Ulrich Drepper, Ahmed Sanaullah, Renato Mancuso, Manos Athanassoulis</i>	
Demonstrating Waffle: A Self-driving Grid Index.....	3954
<i>Dalsu Choi, Hyunsik Yoon, Hyubjin Lee, Yon Dohn Chung</i>	
CM-Explorer: Dissecting Data Ingestion Problems .....	3958
<i>Niels Bylois, Frank Neven, Stijn Vansummeren</i>	
Explaining Differentially Private Query Results With DPXplain .....	3962
<i>Tingyu Wang, Yuchao Tao, Amir Gilad, Ashwin Machanavajjhala, Sudeepa Roy</i>	
Ganos Aero: A Cloud-Native System for Big Raster Data Management and Processing.....	3966
<i>Fei Xiao, Jiong Xie, Zhida Chen, Feifei Li, Zhen Chen, Jianwei Liu, Yinpei Liu</i>	

Demonstration of OpenDBML, a Framework for Democratizing In-Database Machine Learning.....	3970
<i>Mahdi Ghorbani, Amir Shaikhha</i>	
Demonstration of SPARQL-ML: An Interfacing Language for Supporting Graph Machine Learning for RDF Graphs .....	3974
<i>Hussein Abdallah, Waleed Afandi, Essam Mansour</i>	
EQUI-VOCAL Demonstration: Synthesizing Video Queries from User Interactions.....	3978
<i>Enhao Zhang, Maureen Daum, Dong He, Manasi Ganti, Brandon Haynes, Ranjay Krishna, Magdalena Balazinska</i>	
TsQuality: Measuring Time Series Data Quality in Apache IoTDB .....	3982
<i>Yuanhui Qiu, Chenguang Fang, Shaoxu Song, Xiangdong Huang, Chen Wang, Jianmin Wang</i>	
Approximate Queries over Concurrent Updates.....	3986
<i>Congying Wang, Nithin Sastry Tellapuri, Sphoorthi Keshannagari, Dylan Zinsley, Zhuoyue Zhao, Dong Xie</i>	
Solving Hard Variants of Database Schema Matching on Quantum Computers.....	3990
<i>Kristin Fritsch, Stefanie Scherzinger</i>	
Interpretable Clustering of Multivariate Time Series with Time2Feat .....	3994
<i>Angela Bonifati, Francesco Del Buono, Francesco Guerra, Miki Lombardi, Donato Tiano</i>	
DHive: Query Execution Performance Analysis via Dataflow in Apache Hive .....	3998
<i>Chaozu Zhang, Qiaomu Shen, Bo Tang</i>	
MLWHATIF: What If You Could Stop Re-Implementing Your Machine Learning Pipeline Analyses Over and Over? .....	4002
<i>Stefan Grafberger, Shubha Guha, Paul Groth, Sebastian Schelter</i>	
AQUA: Automatic Collaborative Query Processing in Analytical Database .....	4006
<i>Yuchen Peng, Ke Chen, Lidan Shou, Dawei Jiang, Gang Chen</i>	
VisualNeo: Bridging the Gap between Visual Query Interfaces and Graph Query Engines.....	4010
<i>Kai Huang, Houdong Liang, Chongchong Yao, Xi Zhao, Yue Cui, Yao Tian, Ruiyuan Zhang, Xiaofang Zhou</i>	
KG-Roar: Interactive Datalog-based Reasoning on Virtual Knowledge Graphs .....	4014
<i>Luigi Bellomarini, Marco Benedetti, Andrea Gentili, Davide Magnanimi, Emanuel Sallinger</i>	
To UDFs and Beyond: Demonstration of a Fully Decomposed Data Processor for General Data Wrangling Tasks .....	4018
<i>Nico Schäfer, Damjan Gjurovski, Angjela Davitkova, Sebastian Michel</i>	
ChainDash: An Ad-Hoc Blockchain Data Analytics System.....	4022
<i>Yushi Liu, Liwei Yuan, Zhihao Chen, Yekai Yu, Zhao Zhang, Cheqing Jin, Ying Yan</i>	
BrewER: Entity Resolution On-Demand .....	4026
<i>Luca Zecchini, Giovanni Simonini, Sonia Bergamaschi, Felix Naumann</i>	
Visualizing Spreadsheet Formula Graphs Compactly.....	4030
<i>Fanchao Chen, Dixin Tang, Haotian Li, Aditya G. Parameswaran</i>	
DuckPGQ: Bringing SQL/PGQ to DuckDB.....	4034
<i>Daniel Ten Wolde, Gábor Szárnyas, Peter Boncz</i>	

Demo of QueryBooster: Supporting Middleware-based SQL Query Rewriting as a Service.....	4038
<i>Qiushi Bai, Sadeem Alsudais, Chen Li</i>	
Web Connector: A Unified API Wrapper to Simplify Web Data Collection.....	4042
<i>Weiyuan Wu, Pei Wang, Yi Xie, Yejia Liu, George Chow, Jiannan Wang</i>	
FS-Real: A Real-World Cross-Device Federated Learning Platform.....	4046
<i>Dawei Gao, Daoyuan Chen, Zitao Li, Yuexiang Xie, Xuchen Pan, Yaliang Li, Bolin Ding, Jingren Zhou</i>	
ADOps: An Anomaly Detection Pipeline in Structured Logs.....	4050
<i>Xintong Song, Yusen Zhu, Jianfei Wu, Bai Liu, Hongkang Wei</i>	
Portals: A Showcase of Multi-Dataflow Stateful Serverless.....	4054
<i>Jonas Spenger, Chengyang Huang, Philipp Haller, Paris Carbone</i>	
CORNET: Learning Spreadsheet Formatting Rules By Example.....	4058
<i>Mukul Singh, José Cambroner Sanchez, Sumit Gulwani, Vu Le, Carina Negreanu, Gust Verbruggen</i>	
Fanglue: An Interactive System for Decision Rule Crafting.....	4062
<i>Chen Qian, Shiwei Liang, Zhaoyang Wang, Yin Lou</i>	
Odyssey: An Engine Enabling The Time-Series Clustering Journey.....	4066
<i>John Paparrizos, Sai Prasanna Teja Reddy</i>	
ERICA: Query Refinement for Diversity Constraint Satisfaction.....	4070
<i>Jinyang Li, Alon Silberstein, Yuval Moskovitch, Julia Stoyanovich, H. V. Jagadish</i>	
Lingua Manga: A Generic Large Language Model Centric System for Data Curation.....	4074
<i>Zui Chen, Lei Cao, Sam Madden</i>	
XDB in Action: Decentralized Cross-Database Query Processing for Black-Box DBMSes.....	4078
<i>Haralampos Gavriilidis, Leonhard Rose, Joel Ziegler, Kaustubh Beedkar, Jorge-Arnulfo Quiané-Ruiz, Volker Markl</i>	
Interactive Demonstration of EVA.....	4082
<i>Gaurav Tarlok Kakkar, Aryan Rajoria, Myna Prasanna Kalluraya, Ashmita Raju, Jiashen Cao, Kexin Rong, Joy Arulraj</i>	
RESCU-SQL: Oblivious Querying for the Zero Trust Cloud.....	4086
<i>Xiling Li, Gefei Tan, Xiao Wang, Jennie Rogers, Soamar Homs</i>	
DataRinse: Semantic Transforms for Data preparation based on Code Mining.....	4090
<i>Ibrahim Abdelaziz, Julian Dolby, Udayan Khurana, Horst Samulowitz, Kavitha Srinivas</i>	
Demonstrating ADOPT: Adaptively Optimizing Attribute Orders for Worst-Case Optimal Joins via Reinforcement Learning.....	4094
<i>Junxiong Wang, Mitchell Gray, Immanuel Trummer, Ahmet Kara, Dan Olteanu</i>	
Demonstrating GPT-DB: Generating Query-Specific and Customizable Code for SQL Processing with GPT-4.....	4098
<i>Immanuel Trummer</i>	
SHEVA: A Visual Analytics System for Statistical Hypothesis Exploration.....	4102
<i>Vicente Nejar De Almeida, Eduardo Ribeiro, Nassim Bouarour, João Luiz Dihl Comba, Sihem Amer-Yahia</i>	
PikePlace: Generating Intelligence for Marketplace Datasets.....	4106
<i>Shi Qiao, Alekh Jindal</i>	



A Learned Query Rewrite System .....	4110
<i>Xuanhe Zhou, Guoliang Li, Jianming Wu, Jiesi Liu, Zhaoyan Sun, Xinning Zhang</i>	

**Panel**

Will LLMs reshape, supercharge, or kill data science? .....	4114
<i>Alon Halevy, Yejin Choi, Avriella Floratou, Michael J. Franklin, Natasha Noy, Haixun Wang</i>	

**Endowment Awards**

Towards Auto-Generated Data Systems .....	4116
<i>Alvin Cheung, Maaz Bin Safeer Ahmad, Brandon Haynes, Chanwut Kittivorawong, Shadaj Laddad, Xiaoxuan Liu, Chenglong Wang, Cong Yan</i>	

Generations of Knowledge Graphs: The Crazy Ideas and the Business Impact .....	4130
<i>Xin Luna Dong</i>	

The Story of GraphLab – From Scaling Machine Learning to Shaping Graph Systems Research .....	4138
<i>Joseph E. Gonzalez, Yucheng Low</i>	

**Keynotes**

Common Sense: the Dark Matter of Language and Intelligence .....	4139
<i>Yejin Choi</i>	

Modernization of Databases in the Cloud Era: Building Databases that Run Like Legos .....	4140
<i>Feifei Li</i>	

On the Cusp: Computing Thrills and Perils and Professional Awakening .....	4152
<i>Natasa Milic-Frayling</i>	

## PVLDB ORGANIZATION AND REVIEW BOARD - Vol. 16

### Editors in Chief of PVLDB

Georgia Koutrika (Athena Research Center)  
Jun Yang (Duke University)

### Associate Editors of PVLDB

Alkis Simitsis (Athena Research Center)  
Amol Deshpande (University of Maryland at College Park)  
Angela Bonifati (Lyon 1 University)  
Ashwin Machanavajjhala (Duke University/Tumult Labs)  
Badrish Chandramouli (Microsoft Research)  
Boris Glavic (Illinois Institute of Technology)  
Ce Zhang (ETH Zurich)  
Cyrus Shahabi (University of Southern California)  
Dan Olteanu (University of Zurich)  
Eric Lo (The Chinese University of Hong Kong)  
Evaggelia Pitoura (University of Ioannina)  
Evmiria Terzi (Boston University)  
Gustavo Alonso (ETH Zurich)  
Helen (Zi) Huang (University of Queensland)  
Hong Cheng (The Chinese University of Hong Kong)  
Kenneth Ross (Columbia University)  
Khuzaima Daudjee (University of Waterloo)  
Kyuseok Shim (Seoul National University)  
Letizia Tanca (Politecnico di Milano)  
Lucian Popa (IBM Research - Almaden)  
Magdalena Balazinska (University of Washington)  
Meihui Zhang (Beijing Institute of Technology)  
Neoklis Polyzotis (Databricks)  
Nesime Tatbul (Intel Labs and MIT)  
Nikos Mamoulis (University of Ioannina)  
Rachel Pottinger (University of British Columbia)  
Wenjie Zhang (University of New South Wales)  
Wolfgang Gatterbauer (Northeastern University)

Wook-Shin Han (Pohang University of Science and Technology)  
Xiaokui Xiao (National University of Singapore)  
Yannis Velegrakis (University of Trento and Utrecht University)  
Yanyan Shen (Shanghai Jiao Tong University)  
Yi Chen (New Jersey Institute of Technology)  
Yongxin Tong (Beihang University)  
Zhifeng Bao (RMIT University)

### Publication Editors

Manos Athanassoulis (Boston University)  
Kostas Stefanidis (Tampere University)  
Ju Fan (Renmin University of China)

### PVLDB Managing Editor

Wolfgang Lehner (Dresden University of Technology)

### PVLDB Advisory Board

Vanessa Braganholo (Universidade Federal Fluminense)  
Sourav S Bhowmick (Nanyang Technological University)  
Chris Jermaine (Rice University)  
Peter Triantafillou (University of Warwick)  
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)  
Wolfgang Lehner (Dresden University of Technology)  
Felix Naumann (HPI)  
Xuemin Lin (University of New South Wales)  
Georgia Koutrika (Athena Research Center)  
Jun Yang (Duke University)

## Review Board

Abolfazl Asudeh (University of Illinois at Chicago)  
Alexander J Ratner (University of Washington)  
Alexandra Meliou (University of Massachusetts Amherst)  
Amelie Marian (Rutgers University)  
Amir Gilad (Duke University)  
Amir Shaikhha (University of Edinburgh)  
Amrita Roy Chowdhury (University of Wisconsin-Madison)  
Anastasios Kementsietsidis (Google Research)  
Andrew Crotty (Carnegie Mellon University)  
Anna Fariha (Microsoft)  
Anton Dignös (Free University of Bozen-Bolzano)  
Antonios Deligiannakis (Technical University of Crete)  
Arijit Khan (Nanyang Technological University)  
Ashraf Aboulnaga (Qatar Computing Research Institute, HBKU)  
Asterios Katsifodimos (Delft University of Technology)  
Baihua Zheng (Singapore Management University)  
Bin Cui (Peking University)  
Bingsheng He (National University of Singapore)  
Binhang Yuan (ETH Zurich)  
Bogdan Cautis (University Paris-Saclay)  
Bojan Karlas (ETH Zurich)  
Bolin Ding (Data Analytics and Intelligence Lab, Alibaba Group)  
Bolong Zheng (Huazhong University of Science and Technology)  
Bongki Moon (Seoul National University)  
Botong Huang (Alibaba)  
Brad Glasbergen (University of Waterloo)  
Brandon Haynes (Microsoft Gray Systems Lab)  
Cedric Renggli (ETH Zurich)  
Chao Zhang (Lyon 1 University)  
Chen Li (UC Irvine)  
Chengfei Liu (Swinburne University of Technology)  
Chengkai Li (The University of Texas at Arlington)  
Chengliang Chai (Tsinghua University)  
Chong Wang (Amazon)  
Cristian Riveros (PUC Chile)  
Daichi Amagata (Osaka University)  
Dan Kifer (Pennsylvania State University)  
Daniel Deutch (Tel Aviv University)  
Daniel Kang (Stanford University)  
Diego Calvanese (Free University of Bozen)  
Dimitrios Skoutas (Athena Research Center)  
Dimitris Sacharidis (ULB)  
Dirk Habich (TU Dresden)  
Dong Deng (Rutgers University - New Brunswick)  
Dong Wen (University of New South Wales)  
Dong Xie (Pennsylvania State University)  
Dongxiang Zhang (Zhejiang University)  
Dumitrel Loghin (National University of Singapore)  
Elena Ferrari (University of Insubria, Varese)  
Eleni Tzirita Zacharatou (IT University of Copenhagen)  
Essam Mansour (Concordia University)  
Faisal Nawab (University of California at Irvine)  
Fan Zhang (Guangzhou University)  
Fatemeh Nargesian (University of Rochester)  
Fei Chiang (McMaster University)  
Floris Geerts (University of Antwerp)  
Gao Cong (Nanyang Technological University)

George Fakas (Uppsala University)  
George Fletcher (Eindhoven University of Technology the Netherlands)  
George Papadakis (University of Athens)  
George Papastefanatos (Athena Research Center)  
Giovanni Simonini (University of Modena and Reggio Emilia)  
Graham Cormode (University of Warwick)  
Guna Prasaad (Meta Platforms Inc.)  
Guoliang Li (Tsinghua University)  
Guoren Wang (Beijing Institute of Technology)  
Haibo Hu (Hong Kong Polytechnic University)  
Hannes Voigt (Neo4j)  
Haridimos Kondylakis (FORTH-ICS)  
Holger Pirk (Imperial College)  
Huanchen Zhang (Tsinghua University)  
Ibrahim Sabek (MIT)  
Immanuel Trummer (Cornell University)  
Ingo Müller (Google)  
James Cheng (The Chinese University of Hong Kong)  
Jeffrey Xu Yu (The Chinese University of Hong Kong)  
Jens Teubner (TU Dortmund University)  
Jia Yu (Washington State University)  
Jian Lou (Xidian University)  
Jianguo Wang (Purdue University)  
Jianliang Xu (Hong Kong Baptist University)  
Jianxin Li (Deakin University)  
Jiawei Jiang (ETH Zurich)  
Jieming Shi (Hong Kong Polytechnic University)  
Jinfei Liu (Zhejiang University)  
Jing Tang (Hong Kong University of Science and Technology)  
John Liagouris (Boston University)  
John Paparrizos (University of Chicago)  
Joseph Near (University of Vermont)  
Junhao Gan (University of Melbourne)  
K. Selcuk Candan (Arizona State University)  
Kai Wang (University of New South Wales)  
Karima Echiabi (Mohammed VI Polytechnic University)  
Kartik Nayak (Duke University)  
Katja Hose (Aalborg University)  
Kexin Rong (Stanford University)  
Kun Qian (Amazon)  
Kunsoo Park (Seoul National University)  
Kyriakos Mouratidis (Singapore Management University)  
Laks Lakshmanan (University of British Columbia)  
Laurel Orr (Stanford University)  
Lei Cao (MIT)  
Lei Chen (Hong Kong University of Science and Technology)  
Lei Li (Hong Kong University of Science and Technology, Guang Zhou)  
Lijun Chang (The University of Sydney)  
Lin Ma (Carnegie Mellon University)  
Long Yuan (Nanjing University of Science and Technology)  
Lu Qin (UTS)  
Lucas Lersch (Amazon Web Services)  
Lukasz Golab (University of Waterloo)  
Matteo Interlandi (Microsoft)  
Matteo Lissandrini (Aalborg University)  
Matthias Renz (University of Kiel)  
Matthias Weidlich (Humboldt University of Berlin)

Michael Abebe (University of Waterloo)  
 Michael H Boehlen (University of Zurich)  
 Michael Hay (Colgate University/Tumult Labs)  
 Michael Mathioudakis (University of Helsinki)  
 Michal Friedman (ETH)  
 Milos Nikolic (University of Edinburgh)  
 Mirek Riedewald (Northeastern University)  
 Mohamed Sharaf (United Arab Emirates University)  
 Mohammad Sadoghi (University of California, Davis)  
 Mostafa Milani (The University of Western Ontario)  
 Nick Koudas (University of Toronto)  
 Nikolaos Tziavelis (Northeastern University)  
 Nikolay Yakovets (Eindhoven University of Technology)  
 Ning Wang (Beijing Jiaotong University)  
 Oliver A Kennedy (University at Buffalo, SUNY)  
 Panagiotis Bouros (Johannes Gutenberg University Mainz)  
 Panos Kalnis (King Abdullah University of Science and Technology)  
 Panos Vassiliadis (University of Ioannina)  
 Paolo Papotti (EURECOM, France)  
 Periklis Andritsos (University of Toronto)  
 Prashant Pandey (University of Utah)  
 Raghav Kaushik (Microsoft)  
 Rainer Gemulla (University of Mannheim)  
 Raul Castro Fernandez (University of Chicago)  
 Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)  
 Renata Borovica-Gajic (University of Melbourne)  
 Reynold Cheng (The University of Hong Kong, China)  
 Riccardo Torlone (Roma Tre University)  
 Ronghua Li (Beijing Institute of Technology)  
 Ryan C Marcus (MIT)  
 Ryan Stutsman (University of Utah)  
 Sai Wu (Zhejiang Univ)  
 Sairam Gurajada (Apple)  
 Sebastian Link (University of Auckland)  
 Senjuti Basu Roy (New Jersey Institute of Technology)  
 Seokki Lee (University of Cincinnati)  
 Shantanu Sharma (New Jersey Institute of Technology)  
 Shaoxu Song (Tsinghua University)  
 Shiyu Yang (Guangzhou University)  
 Shuai Ma (Beihang University)  
 Sibo Wang (The Chinese University of Hong Kong)  
 Siqiang Luo (Nanyang Technological University)  
 Sourav S Bhowmick (Nanyang Technological University)  
 Spyros Blanas (The Ohio State University)  
 Srikanta Bedathur (IIT Delhi)  
 Stefania Dumbrava (ENSIIE)  
 Stefano Paraboschi (Universita' degli Studi di Bergamo)  
 Steffen Zeuch (DFKI Berlin)  
 Steven E Whang (KAIST)  
 Stijn Vansummeren (Hasselt University)  
 Sudipto Das (Amazon Web Services)  
 Tarique Siddiqui (Microsoft Research)  
 Theodore Dalamagas (Athena Research Center)  
 Thomas Neumann (TU Munich)  
 Tian Li (Carnegie Mellon University)  
 Tianhao Wang (University of Virginia)  
 Tianzheng Wang (Simon Fraser University)  
 Tien Tuan Anh Dinh (Singapore University of Technology and Design)  
 Torben Bach Pedersen (Aalborg University)  
 Utku Sirin (Harvard University)  
 Vasiliki Kalavri (Boston University)  
 Vassilios S Verykios (Hellenic Open University)  
 Walid G Aref (Purdue)  
 Wang-Chiew Tan (Facebook AI)  
 Weiguo Zheng (Fudan University)  
 Wendy Hui Wang (Stevens Institute of Technology)  
 Wentao Wu (Microsoft Research)  
 Xi He (University of Waterloo)  
 Xiang Lian (Kent State University)  
 Xiangmin Zhou (RMIT University)  
 Xiangyao Yu (University of Wisconsin-Madison)  
 Xiaochun Yang (Northeastern University)  
 Xiaofei Zhang (University of Memphis)  
 Xiaoyang Wang (University of New South Wales)  
 Xin Cao (University of New South Wales)  
 Xin Huang (Hong Kong Baptist University)  
 Yan Zhao (Aalborg University)  
 Yang Cao (Kyoto University)  
 Yao Lu (Microsoft Research)  
 Ye Yuan (Beijing Institute of Technology)  
 Yeye He (Microsoft Research)  
 Yi Yu (NII)  
 Yinghui Wu (Case Western Reserve University)  
 Yingxia Shao (BUPT)  
 Yixiang Fang (The Chinese University of Hong Kong, Shenzhen)  
 Yongluan Zhou (University of Copenhagen)  
 You Peng (University of New South Wales)  
 You Wu (Google)  
 Yufei Tao (The Chinese University of Hong Kong)  
 Yuncheng Wu (National University of Singapore)  
 Yuyu Luo (Tsinghua University)  
 Zeke Wang (Zhejiang University)  
 Zhiwei Zhang (Beijing Institute of Technology)  
 Zhongle Xie (Zhejiang University)  
 Zhuoyue Zhao (University at Buffalo - SUNY)  
 Ziawasch Abedjan (Leibniz University Hannover)  
 Zimu Zhou (Singapore Management University)

## INDUSTRIAL TRACK CHAIRS AND REVIEWERS - Vol. 16

### Industrial Track Program Chairs

Abdul Quamar (Google)

Yuanyan Tian (Microsoft Gray Systems Lab)

### Industrial Track Program Committee Members

Amol Deshpande (University of Maryland at College Park)

Beng Chin Ooi (NUS)

Chuan Lei (Amazon Web Services)

Danica Porobic (Oracle)

Divy Agrawal (University of California, Santa Barbara)

Eser Kandogan (Megagon Labs)

Fabian Panse (Universität Hamburg)

Fatma Ozcan (Google)

Feifei Li (Alibaba Group)

Gustavo Alonso (ETHZ)

Hanuma Kodavalla (Microsoft)

Ihab Ilyas (U. of Waterloo and Apple)

Jagan Sankaranarayanan (Google Inc)

Karthik Ramachandra (Microsoft Azure SQL India)

Lyublena Antova (Meta)

Matthias Boehm (Technische Universität Berlin)

Mohamed Soliman (Apple)

Norman May (SAP SE)

Oktie Hassanzadeh (IBM Research)

Rana Alotaibi (Microsoft Gray Systems Lab)

Roger Barga (Amazon)

Shaleen Deep (Microsoft Gray Systems Lab)

Umesh Dayal (Hitachi America)

Venkata Vamsikrishna Meduri (IBM Research, Almaden)

Wolfram Wingerath (University of Oldenburg)

Xiangyao You (University of Wisconsin-Madison)

Yannis Katsis (IBM Research)

Yingjun Wu (RisingWave Labs)

Chris Douglas (UC Berkeley)

Chen Luo (Snowflake)

Liang Zhang (TigerGraph)

Chang Ge (University of Minnesota)

Bailu Ding (Microsoft Research)

Juang Colmenares (LinkedIn)

Yannis Chronis (Google)

Vijayshankar Raman (Google)

Carlo Curino (Microsoft)

Yiwen Zhu (Gray Systems Lab)

Lukas Rupprecht (Databricks)

Fotis Psallidas (Gray Systems Lab)

## DEMONSTRATION TRACK CHAIRS AND REVIEWERS - Vol. 16

### Demonstration Track Program Chairs

Alekh Jindal (Microsoft Research)  
Carsten Binnig (TU Darmstadt)  
Jennie Rogers (Northwestern University)

### Demonstration Track Program Committee Members

Alexander van Renen (Friedrich-Alexander-Universität Erlangen-Nürnberg)  
Amir Gilad (Duke University)  
Amit Somech (Bar-Ilan University)  
Andrew Crotty (Carnegie Mellon University)  
Avigdor Gal (Technion -- Israel Institute of Technology)  
Behrooz Omidvar-Tehrani (Amazon)  
Bin Cui (Peking University)  
Cong Yan (Microsoft research)  
El Kindi Rezig (MIT)  
Enzo Veltri (University of Basilicata)  
Gabriel Ghinita (Hamad Bin Khalifa University)  
Georgia Troullinou (FORTH-ICS)  
Harish Doraiswamy (Microsoft Research India)  
Hiroaki Shiokawa (University of Tsukuba)  
Jaroslaw Szlichta (York University and IBM CAS)  
Jialin Ding (AWS)  
Jiangwei Zhang (National University of Singapore)  
Ju Fan (Renmin University of China)  
Jyoti Leeka (Microsoft)  
Krishna Kantikiran Pasupuleti (Oracle)  
Leonard Ribeiro (Federal University of Goiás)  
Lihong He (IBM Almaden Research Center)  
Manisha Luthra (TU Darmstadt)  
Mayuresh Kunjir (Amazon AWS)  
Milos Nikolic (University of Edinburgh)  
Mourad Ouzzani (Qatar Computing Research Institute, HBKU)  
Muhammad El-Hindi (TU Darmstadt)  
Nan Tang (Qatar Computing Research Institute, HBKU)  
Nantia Makrynioti (CWI and RelationalAI)  
Ramon Lawrence (UBC)  
Rebecca Taft (Cockroach Labs)  
Renata Borovica-Gajic (University of Melbourne)  
Roe Shraga (Technion - Israel Institute of Technology)  
Rong Zhu (Alibaba Group)  
Sainyam Galhotra (University of Chicago)  
Sang-Wook Kim (Hanyang University, Korea)  
Simon Razniewski (Max-Planck-Institut für Informatik, Germany)  
Subarna Chatterjee (Harvard University, USA)  
Subhadeep Sarkar (Boston University)  
Tiago de Melo (Universidade do Estado do Amazonas)  
Timo Bang (TU Darmstadt)  
Tsz Nam (Edison) Chan (Hong Kong Baptist University)  
Varun Pandey (TU Berlin)  
Venkatesh Emani (Microsoft)  
Verena Kantere (National Technical University of Athens, Greece)  
Y.C. Tay (National University of Singapore)  
Yaron Kanza (AT&T Labs-Research)  
Yingxia Shao (BUPT)  
Yiru Chen (Columbia University)  
Yiwen Zhu (Microsoft)  
Yuval Moskovitch (Ben Gurion University)  
Zainab Abbas (KTH Royal Institute of Technology)  
Zeyuan Shang (Einblick Analytics)  
Zhi Yang (Peking University)

## TUTORIAL TRACK CHAIRS AND REVIEWERS - Vol. 16

### Tutorial Track Program Chairs

Senjuti Basu Roy (New Jersey Institute of Technology)

Steven E. Whang (KAIST)

### Tutorial Track Program Committee Members

Chao Zhang (Tsinghua University)

Chenhao Ma (Chinese University of Hong Kong)

Chuan Xiao (Osaka University and Nagoya University)

Jayant Haritsa (Indian Institute of Science)

Jelle Hellings (McMaster University)

Jianbin Qin (Shenzhen University)

Karima Echihabi (Mohammed VI Polytechnic University)

Mareike Schmidt (Universität Hamburg)

Ramon Antonio Rodrigues Zalipynis (HSE University)

Simon Razniewski (Bosch Center for AI)

Suyash Gupta (UC Berkeley)

Wasay Abdul (Intel Labs)

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

Zoi Kaoudi (TU Berlin)

## LETTER FROM THE EDITORS IN CHIEF

It is our pleasure to present the twelfth issue of PVLDB (Proceedings of the VLDB) Volume 16. While the first eleven issues covered research papers accepted to PVLDB and presented at the VLDB 2023 Conference, Issue 12 includes the keynote talks, panel, workshops, as well as peer-reviewed industrial papers, demonstrations, and tutorials, which were also part of the VLDB 2023 program.

The Industrial, Applications, and Experience Track of VLDB 2023 covers all aspects of innovative commercial or industrial-strength data management systems and solutions, including novel real-world applications of data management systems and experience in applying recent research advances to real-world problems. This year, we received 76 submissions, out of which we accepted 29 papers (38%) after a review by the Industrial Track Program Committee. Each paper received 3 reviews. The Industrial Program Chairs in coordination with the PC Chairs made the final decisions after calibration. The authors and reviewers were required to declare their COIs, and the paper assignments to reviewers were made to ensure that there were no COIs. For the final decision, PC Chairs stepped in to handle submissions that had COIs with the Industry Program Chairs. The accepted papers covered a wide range of real-world data management solutions including machine learning for systems, systems for machine learning, benchmarking, performance and resource optimization for cloud-based systems, governance, as well as novel database systems for real-world use cases.

The Demonstration Track of VLDB 2023 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 134 submissions, which is similar to the previous years. After a rigorous review process, 53 (39.5%) of these submissions were accepted. Each demonstration proposal was reviewed by at least three reviewers; some had four. 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 VLDB 2023 covers state-of-the-art research, development, and applications in data management or related areas, including interdisciplinary directions. This year, we received 25 submissions, out of which we accepted 9 short (1.5 hours) and 2 long (3 hours) tutorials, representing a 44% acceptance rate. All accepted tutorials were lecture-style, some of which have been recommended to include some demonstration components. The Tutorial Program Committee 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 on how many proposals to accept. The accepted proposals cover a wide range of topics including deep learning, natural language interfaces, visualization, user-defined functions, modern networks for database systems, graph processing, data and AI model markets, collaborative data analytics, private information retrieval, and time series data mining.

VLDB 2023 received 15 workshop proposals. After reviewing the proposals and evaluating their merits based on their history, relevance of topics, and distinctiveness, we selected 13 high-quality workshops (5 full-day and 8 half-day). Similar to previous years, we erred on the side of accepting a larger number of workshops in order to ensure a good diversity of topics and allow new and revived workshops to establish themselves with sufficiently many submissions. The following workshops were accepted as full-day:

- ADMS: 14th International Workshop on Accelerating Analytics and Data Management Systems Using Modern Processor and Storage Architectures; Rajesh Bordawekar, Tirthankar Lahiri (August 28)
- CDMS: 2nd International Workshop on Composable Data Management Systems 2023; Satyanarayana R Valluri, Mohamed Zait (August 28)
- CloudDB: 1st Workshop on Cloud Databases; Sanjay Krishnan, Guoliang Li, Jiannan Wang, Kai Zeng (August 28)
- QDB: 12th International Workshop on Quality in Databases; Lisa Ehrlinger, Hazar Harmouch Ihab Ilyas Felix Naumann (August 28)
- TPCTC: 15th TPC Technology Conference on Performance Evaluation and Benchmarking 2023; Raghunath Nambiar, Meikel Poess (August 28)

The following are half-day workshops:

- AIDB: 5th Applied AI for Database Systems and Applications; Umar Farooq Minhas, Yingjun Wu (September 1)
- DEco: 2nd International Workshop on Data Ecosystem; Sandra Geisler, Maria-Esther Vidal Cinzia Cappiello (August 28)
- LLM: 1st Workshop on Databases and Large Language Models; Haixun Wang, Alon Halevy (September 1)
- Poly: 6th Polystore systems for heterogeneous data in multiple databases with privacy and security assurances, El Kindi Rezig, Vijay Gadepally Michael Stonebraker (September 1)



- QDSM: 1st International Workshop on Quantum Data Science and Management; Sven Groppe, Jiaheng Lu, Wolfgang Mauerer, Le Gruenwald (September 1)
- SDA: 1st Workshop on Serverless Data Analytics; Tamer Özsu, Xun Xue (September 1)
- SDM: 1st (Mini) Symposium on Data Markets; Jian Pei, Raul Castro Fernandez Xiaohui Yu (August 28)
- TaDA: 1st Workshop on Tabular Data Analysis; Vasilis Efthymiou, Sainyam Galhotra Oktie Hassanzadeh Ernesto Jiménez-Ruiz Kavitha Srinivas (September 1)

The PhD Workshop features 12 student presentations, a keynote by Wang-Chiew Tan (Facebook AI) and a panel on “Good Reviewing Habits.”

VLDB 2023 features three keynotes:

- Common Sense: the Dark Matter of Language and Intelligence: Dr. Yejin Choi, Wissner-Slivka Professor, Paul G. Allen School of Computer Science & Engineering, University of Washington
- Modernization of Databases in the Cloud Era: Building Databases that Run Like Legos: Dr. Feifei Li, Vice President, Alibaba Group; Director of Alibaba Cloud Database Product Business Unit; Director of DAMO Academy Database Lab
- On the Cusp: Computing Thrills and Perils and Professional Awakening: Prof. Dr. Natasa Milic-Frayling, CEO, Intact Digital Ltd.; Research Director, Arabic Language Technologies, Qatar Computing Research Institute, Doha, Qatar; Professor Emerita, University of Nottingham, UK

VLDB 2023 also features one panel on “Will LLMs Reshape, Supercharge, or Kill Data Science?” moderated by Dr. Alon Halevy (Meta Reality Labs), with panelists Yejin Choi (University of Washington and AI2), Avriella Floratou (Microsoft), Michael J. Franklin (University of Chicago), Natasha Noy (Google), and Haixun Wang (Instacart).

Besides the Best Paper and Demonstration awards announced at the VLDB 2023 Conference, three VLDB Endowment Awards are made:

- 2023 VLDB Early Career Research Contribution Award: Prof. Alvin Cheung (UC Berkeley)
- 2023 VLDB Women in Database Research Award: Dr. Xin Luna Dong (Meta)
- 2023 VLDB Test of Time Award: Yucheng Low (XetHub Inc.), Joseph Gonzalez (UC Berkeley), Aapo Kyrola (Missing Exponent Oy), Danny Bickson (Visual Layer), Carlos Guestrin (Stanford University), Joseph M. Hellerstein (UC Berkeley) for their VLDB 2012 paper “Distributed Graphlab: A Framework for Machine Learning in the Cloud”

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 16.

Georgia Koutrika and Jun Yang  
Editors-in-Chief of PVLDB Vol. 16  
Program Chairs for VLDB 2023

Abdul Quamar and Yuanyan Tian  
Industrial Program Chairs

Alekh Jindal, Carsten Binnig, and Jennie Rogers  
Demonstration Program Chairs

Senjuti Basu Roy and Steven E. Whang  
Tutorial Program Chairs

Matthias Boehm and Aaron Elmore  
Workshop Chairs

Vasilis Efthymiou and Xiao Hu  
PhD Workshop Chairs

Xiaofang Zhou and Alan Fekete  
Panel Chairs