Proceedings of the VLDB Endowment

Volume 18, No. 4 – December 2024

PVLDB – Proceedings of the VLDB Endowment

Volume 18, No. 4, December 2024.

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

Volume 18, Number 4, December 2024

Pages i – viii and 929-1263

ISSN 2150-8097

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

TABLE OF CONTENTS

Front Matter

Research Papers

PVLDB ORGANIZATION AND REVIEW BOARD - Vol. 18

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

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

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

PVLDB Managing Editors
Wolfgang Lehner (TU Dresden)

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

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

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

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

LETTER FROM THE EDITORS IN CHIEF

This fourth issue of PVLDB Volume 18, the last one for the year 2024, presents 25 papers from three categories: Scalable Data Science (2), Experiment, Analysis & Benchmark (4), and Regular Research (19). The two most popular primary subject areas for these papers are: Machine Learning, AI, and Databases; Data Mining and Analytics.

In Volume 18, we are pleased to see an elevated interest in submissions to the Experiment, Analysis & Benchmark paper category (about a 15% increase compared to those in the previous volume).

The Experiment, Analysis & Benchmark (EA&B) paper category of PVLDB has been serving the data management community by offering insights on new and established problems alike, helping forge new methodologies for evaluating research, and uncovering opportunities for future research. These papers offer comprehensive studies that evaluate existing algorithms, systems, and designs under various workloads to provide valuable insights into performance, scalability, and resource utilization. We also welcome submissions that propose new benchmarks that can help conduct a more streamlined evaluation of new approaches. A common fundamental element of all EA&B papers is artifact availability and reproducibility. We view this as a key requirement to ensure the maximal impact of the analysis presented and benchmarks proposed.

The special nature of EA&B papers also makes the review process unique. We ask the reviewers to shift their focus from evaluating the innovation in terms of a new design, system, or algorithm to evaluating the depth of the analysis and the quality of the observations of the analysis. As mentioned above, artifact availability and full result reproducibility are key aspects of the review process, along with the new research directions and opportunities they uncover. In this PVLDB volume, we decided to have dedicated Associate Editors for the EA&B category, which has helped maintain a consistent mindset throughout the entire review process, and ensure high-quality reviews for this type of submissions.

The first four issues of PVLDB Volume 18 published so far have featured 8 EA&B papers on various topics (4 of them appearing in this issue) with a strong focus on Graph Neural Networks. These include performance comparisons of different training approaches (full-graph vs. mini-batch), high-level observations after long training sessions, and the impact of graph re-ordering. Other topics include joins on LSM-based storage engines with RocksDB as the testing platform, densest subgraph discovery, an experimental survey of GPU-accelerated dynamic graph management, reliability analysis of stream processing systems, and a new benchmark of scalable graph clustering.

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

Manos Athanassoulis, Ioana Manolescu, Beng Chin Ooi
Associate Editors for the Experiment, Analysis & Benchmark (EA&B) Paper Category

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