Proceedings of the VLDB Endowment

Volume 14, No. 6 – February 2021

PVLDB – Proceedings of the VLDB Endowment

Volume 14, No. 6, February 2021.

All papers published in this issue will be presented at the 47th International Conference on Very Large Data Bases, Copenhagen, Denmark, 2021.

Copyright 2021 VLDB Endowment

Volume 14, Number 6, February 2021

Pages i – vii and 863 - 1123

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. 14

Review Board

Abolfazl Asudeh (University of Illinois)
Ahmed Eldawy (University of California, Riverside)
Alan Fekete (University of Sydney)
Alekh Jindal (Microsoft)
Alex Ratner (University of Washington)
Altigran da Silva (Universidade Federal do Amazonas)
Anthony Tung (National University of Singapore)
Antonios Deligiannakis (Technical University of Crete)
Arijit Khan Nanyang (Technological University, Singapore)
Arnau Prat (Sparsity Technologies)
Ashwin Machanavajjhala (Duke University)
Asterios Katsifodimos (Technical University of Delft)
Avrilia Floratou (Microsoft)
Babak Salimi (University of Washington)
Badrish Chandramouli (Microsoft Research)
Beng Chin Ooi (National University of Singapore)
Bin Yang (Aalborg University)
Boris Glavic (Illinois Institute of Technology)
Byron Choi (Hong Kong Baptist University)
Carlos Scheidegger (University of Arizona)
Carsten Binnig (Technical University of Darmstadt)
Ce Zhang (ETH Zurich)
Chengfei Liu (Swinburne University of Technology)
Chengkai Li (University of Texas at Arlington)
Chris Jermaine (Rice University)
Christian Bizer (University of Mannheim)
Cong Yu (Google)
Daisy Zhe Wang (University of Florida)
Danica Porobic (Oracle)
Davide Mottin (Aarhus University)
Dimitris Papadias (Hong Kong University of Science and Technology)
Dong Deng (Rutgers University)
Eric Lo (Chinese University of Hong Kong)
Essam Mansour (Concordia University)
Fatma Ozcan (IBM Research)
Flip Korn (Google)
Florin Rusu (University of California, Merced)
Fotis Psallidas (Microsoft)
Francesco Bonchi (ISI Foundation)
Gao Cong (Nanyang Technological University)
George Fletcher (Technical University of Eindhoven)
Georgia Koutrika (Athena Research Center)
Hao Wei (Amazon)
Heiko Mueller (New York University)
Hong Cheng (Chinese University of Hong Kong)
Hongzhi Wang (Harbin Institute of Technology)
Hung Ngo (RelationalAI)
Immanuel Trummer (Cornell University)
Ingo Müller (ETH Zürich)
Jana Giceva (Technical University of Munich)
Jennie Rogers (Northwestern University)
Jeong-Hyon Hwang (University at Albany, State University of New York)
Jiaheng Lu (University of Helsinki)
Jianliang Xu (Hong Kong Baptist University)
Jianxin Li (Deakin University)
Jignesh Patel (University of Wisconsin)
Johann Gamper (Free University of Bozen-Bolzano)
Johannes Gehrke (Microsoft)
Jonas Traub (Technical University of Berlin)
Joy Arulraj (Georgia Tech)
Ju Fan (Renmin University of China)
K. Selçuk Candan (Arizona State University)
Kai Zeng (Alibaba)
Katja Hose (Aalborg University)
Ken Salem (University of Waterloo)
Kenneth A. Ross (Columbia University)
Khuzaima Daudjee (University of Waterloo)
Konstantinos Karanasos (Microsoft)
Laurel Orr (Stanford University)
Lei Chen (Hong Kong University of Science and Technology)
Lei Zou (Peking University)
Li Xiong (Emory University)
Lu Chen (Aalborg University)
Lu Qin (University of Technology Sydney)
Manasi Vartak (Verta)
Manos Athanassoulis (Boston University)
Manos Karpathiotakis (Facebook)
Marco Serafini (University of Massachusetts Amherst)
Marcos Antonio Vaz Salles (University of Copenhagen)
Mark Callaghan (MongoDB)
Markus Weimer (Microsoft)
Matei Zaharia (Stanford University, Databricks)
Matteo Interlandi (Microsoft)
Matthaios Olma (Microsoft Research)
Meihui Zhang Beijing (Institute of Technology)
Miao Qiao (University of Auckland)
Michael H. Böhlen (University of Zurich)
Michael Cafarella (University of Michigan)
Mirek Riedewald (Northeastern University)
Mohamed Mokbel (Qatar Computing Research Institute)
Mohamed Sarwat (Arizona State University)
Mohammad Sadoghi (University of California, Davis)
Mourad Ouzzani (Qatar Computing Research Institute, Hamad Bin Khalifa University)
Muhammad Aamir Cheema (Monash University)
Murat Demirbas (University at Buffalo, SUNY)
Nan Tang (Qatar Computing Research Institute, Hamad Bin Khalifa University)
Nick Koudas (University of Toronto)
Nikolaus Augsten (University of Salzburg)
Norman May (SAP)
Norman Paton (University of Manchester)
Odysseas Papapetrou (Technical University of Eindhoven)
Oliver A. Kennedy (University at Buffalo, SUNY)
Paolo Merialdo (Roma Tre University)
Paraschos Koutris (University of Wisconsin – Madison)
Peter Boncz (Centrum Wiskunde & Informatica)
Qin Zhang Indiana (University Bloomington)
Raja Appuswamy (Eurecom)
Ralf Schenkel (University of Trier)

Raul Castro Fernandez (University of Chicago)
Raymond Chi-Wing Wong (Hong Kong University of Science and Technology)
Reynold Cheng (The University of Hong Kong)
Reza Akbarinia (INRIA)
Ruoming Jin (Kent State University)
Ryan Johnson (Amazon Web Services)
S. Sudarshan (IIT Bombay)
Sanjay Krishnan (University of Chicago)
Saravanan Thirumuruganathan (Qatar Computing Research Institute, Hamad Bin Khalifa University)
Sebastian Schelter (University of Amsterdam)
Semih Salihoglu (University of Waterloo)
Senjuti Basu Roy (New Jersey Institute of Technology)
Shaoxu Song (Tsinghua University)
Shimin Chen (Chinese Academy of Sciences)
Sibo Wang (The Chinese University of Hong Kong)
Silu Huang (Microsoft Research)
Spyros Blanas (Ohio State University)
Srikanth Kandula (Microsoft Research)
Steffen Zeuch (German Research Centre for Artificial Intelligence - DFKI)
Stijn Vansummeren (Université libre de Bruxelles)
Sudeepa Roy (Duke University)
Sudip Roy (Google)
Tamer Özsu (University of Waterloo)
Themis Palpanas (University of Paris, French University Institute - IUF)
Tianzheng Wang (Simon Fraser University)
Tingjian Ge (University of Massachusetts, Lowell)
Thomas Heinis (Imperial College)
Thomas Neumann (Technical University of Munich)
Toon Calders (Universiteit Antwerpen)
Umar Farooq Minhas (Microsoft Research)
Viktor Leis (Friedrich Schiller University Jena)
Walid Aref (Purdue University)
Wei-Shinn Ku (Auburn University)
Weiren Yu (University of Warwick)
Wendy Hui Wang (Stevens Institute of Technology)
Wenjie Zhang (University of New South Wales)
Wolfgang Gatterbauer (Northeastern University)
Xi He (University of Waterloo)
Xiang Zhao (National University of Defence Technology)
Xiangyao Yu (University of Wisconsin – Madison)
Xiaokui Xiao (National University of Singapore)
Xiaolan Wang (Megagon Labs)
Xin Cao (University of New South Wales)
Xu Chu (Georgia Tech)
Yannis Velegrakis (Utrecht University)
Ye Yuan (Beijing Institute of Technology)
Yeye He (Microsoft Research)
Ying Zhang (University of Technology Sydney)
Yinghui Wu (Case Western Reserve University)
Yongjoo Park (University of Illinois at Urbana-Champaign)
Yongxin Tong (Beihang University)
Yu Yang (City University of Hong Kong)
Yuchen Li (Singapore Management University)
Yudian Zheng (Twitter)
Yunjun Gao (Zhejiang University)
Zechao Shang (University of Chicago)
Zhenjie Zhang (Singapore R&D, Yitu Technology Ltd.)
Zhewei Wei (Renmin University of China)
Ziawasch Abedjan (Technical University of Berlin)
Zoi Kaoudi (Technical University of Berlin)

EDITORIAL

The sixth issue of PVLDB's Volume 14, which we have the pleasure to announce with this letter, covers a variety of exciting topics in the areas of core database systems, machine learning, graph data, data streams, and data accessibility. We were delighted to see that the 21 papers of this issue, which represent the research effort of our community, show a strong commitment to real-world problems in highly relevant areas, such as medicine, green-IT and digitalization: Never has it been more important to produce accurate results, be careful with our limited resources and build reliable and safe systems.

This month, 17 regular research papers, three scalable data science papers, and one vision paper have been published. With eight research projects in the area of graph data, graphs clearly dominate the topic list of the current issue: Leo and Boncz propose a novel analytical database for structural dynamic graphs, Tsitsulin et al. found a novel graph embedding, Yang et al. accelerate the counting of bi-triangles in bipartite graphs, Zhu et al. search for multiple targets in graphs, Fujiwara et al. improve anchor graph hashing, Tsamoura et al. accelerate the materialization of knowledge bases with trigger graphs, Gao et al. facilitate graph neural networks for community discovery, and Li et al. carefully study motif-paths. Similar to our past five issues, research on and with machine learning techniques also plays a major role in Issue 6: Yin et al. make CPUs competitive for deep learning tasks, Liu et al. study model marketplaces, Tata et al. use machine learning to extract structured information from form-like documents, and Gao et al. facilitate graph neural networks for the discovery of communities that contain a certain query vertex. Issue 6 also features a solid set of core database research papers: Thorne et al. envision neural databases, Gubner and Boncz explore the query engine design space with a novel framework, Sun et al. demonstrate how to improve the throughput of encrypted databases, Chen and Nguyen support DNA databases in finding k-similar substring DNAs to any given query pattern, and Zhang et al. introduce database-style query optimization techniques into a programming language compiler. In the area of data streaming, Wang et al. improve the accuracy of estimating flow spread in high-rate data streams, Walter Cai et al. optimize threshold functions over streams, and Wang et al. improve the performance of monitoring time series. Last but not least, Rahman et al. feature dynamic hierarchical overviews to make large spreadsheets accessible and Macke et al. improve computational notebooks so that the notebooks can prevent erroneous interactions.

Volume 14 as a whole also marks a special milestone in the history of PVLDB, because we introduced several changes to the publication format in the background. All published papers are now PDF/A-compliant and, hence, archivable. In this way, PVLDB makes an effort towards long-term preservation of its content, which directly ties in with its promise to authors and readers that the research results are kept accessible. In our role as proceedings chairs, we also retired the old PVLDB style template in which many errors had accumulated over the years. Submissions to PVLDB are now based on the ACM style template, which is well maintained and updated regularly. With the new template, authors can switch their papers more easily between ACM conferences and benefit from a modern design. To ease the process of starting a new publication, we now provide a PVLDB template project on Overleaf - a collaborative cloud-based LaTeX editor. Another feature of the new style template is a well-defined artifact availability tag that authors can activate by providing a link to their project artifacts, which primarily include code and data, but may also be scripts, measurements and other assets. The linked artifacts are checked in the publication process and contribute to PVLDB's reproducibility efforts. The community feedback to all these changes was largely positive and sparked many suggestions for further improvements, such as the use of certain LaTeX features in the review process and a stricter template enforcement – features that we might see in PVLDB Volume 15.

As proceedings chairs, we want to take the opportunity of writing this editorial to thank all authors for their help in producing high-quality camera-ready copies of their papers. Fitting all content on 12 pages and still adhering to the format regulations is often not an easy task. In particular cases, last-minute actions and fast responses were needed to publish an issue in time, for which we are very grateful.

We hope you all enjoy Issue 6 and that it may inspire new ideas.

Hannes Mühleisen and Thorsten Papenbrock
Publication Editors of PVLDB Volume 14 & Proceedings Chairs for VLDB 2021