**GUEST EDITORIAL**

# Special issue on best papers of VLDB 2017

Peter Boncz[1] · Kenneth Salem[2]

VLDB 2017 was held in Munich, Germany, from August 28, 2017, through September 1, 2017. The research paper track received 774 submissions, of which about 18% were ultimately accepted to appear at the conference. Based on the results of the reviewing process and nominations from the program committee, we selected a short list of five candidates for the conference best paper award. A best paper committee, consisting of Alan Fekete, Mei Hsu, and S. Sudarsan, reviewed the shortlisted papers to select a best paper for the conference. In addition, the authors of the five candidate papers were invited to submit extended versions of their work for consideration for this special issue. Four extended papers were submitted, each of which went through the Journal's peer review process. These papers nicely represent the breadth of work that appeared at the 2017 VLDB conference.

In their paper "Explaining Natural Language Query Results," authors Daniel Deutch, Nave Frost, and Amir Gilad consider the problem of providing provenance to non-expert database users, to help them understand the results of their queries. In particular, they consider how to provide explanations, in natural language, of the answers to database queries that are posed in natural language. As was noted by the best paper committee, "two key insights of the paper are to generate answers by leveraging the structure from the original natural language queries, and to provide simpler explanations by factorizing and summarizing provenance." The paper shows that such explanations can be generated quickly and effectively.

In "ORPHEUSDB: Bolt-On Versioning for Relational Databases," Silu Huang, Liqi Xu, Jialin Liu, Aaron Elmore, and Aditya Parameswaran focus on the problem of supporting dataset versioning. Dataset versioning occurs as a result of iterative data analysis, during which many versions of the same dataset are produced. Data scientists need to be able to track the versions and query across them. This paper considers the problem of how to compactly represent versions in a relational database, while effectively supporting efficient version retrieval. The paper develops a fast and effective algorithm, called LYRESPLIT, that can be used to choose a representation and maintain it as new versions are added.

The paper "ENTROPYDB: Probabilistic Approach to Approximate Query Processing," by Laurel Orr, Magda Balazinska, and Dan Suciu, describes a new approach to use probabilistic models to generate small, query-able summaries for interactive data exploration. The key technical innovation is that the model admits a compact polynomial representation. A naive approach would scale with the number of possible tuples, thus exponentially with attribute domain size. This paper also presents an optimized implementation that uses bitmaps and caching to build and evaluate the polynomial.

Finally, in "Adaptive Partitioning and Indexing for in situ Query Processing," Matthaios Olma, Manos Karpathiotakis, Ioannis Alagiannis, Manos Athanassoulis, and Anastasia Ailamaki present a new online partitioning and indexing scheme, along with a partitioning and indexing tuner tailored for in situ querying engines. This design improves query execution time by taking into account user query patterns, to partition raw data files logically and build for each partition lightweight partition-specific indexes. The authors also built an in situ query engine called SLALOM, which employs this adaptive partitioning and builds non-obtrusive indexes in different partitions on-the-fly based on lightweight query access pattern monitoring. As a result of its lightweight nature, SLALOM is demonstrated to achieve efficient query processing over raw data minimal memory consumption.

We hope that you enjoy these examples of the great work that appeared at VLDB 2017. We are grateful to the VLDB conference and VLDB Journal reviewers for their work in assessing and guiding the papers and to the VLDB best paper committee for its care and diligence. Finally, we thank

✉ Peter Boncz
  Peter.Boncz@cwi.nl

  Kenneth Salem
  ken.salem@uwaterloo.ca

[1] CWI, Amsterdam, The Netherlands

[2] University of Waterloo, Waterloo, Canada

Ⓐ Springer

the authors for the substantial work they have put into the extended papers that appear in this issue.

Peter Boncz and Kenneth Salem
Guest Editors of the Special Issue

Program Chairs of VLDB 2017
Editors in Chief of Vol 10. of PLVDB