# Editor's Notes

Welcome to the December 2014 issue of the ACM SIGMOD Record!

The issue opens with a Database Principles article by Amsterdamer and Milo on "Foundations of Crowd Data Sourcing." Crowdsourcing has been a popular approach to harnessing data and processing data in domains where human contributions to these tasks are highly valuable, such as websites of reviews and ratings. This article provides a timely survey of important recent work that offers sound scientific foundations for crowd-sourcing. More specifically, it addresses the challenges in modeling crowd sourcing for two main tasks: (1) harvesting data, for which a crowd mining approach is proposed to identify statistically significant patterns within the habits and preferences of the crowd, and analyzed with complexity results; (2) processing data with the help of the crowd, for which classic database operators, including top-$k$ and group-by, are analyzed with theoretical results in this unique problem setting.

The Research and Vision Articles Column features two articles. The first one, by Zliobaite et al, presents a vision of hardware-driven design of low-energy data analysis algorithms. It begins by surveying the recent trend in hardware development, pointing to the need of designing algorithms that minimize communication (memory accesses and data transfer between cores) and related energy consumption. It then considers a range of data mining algorithms as example data analysis tasks, and proposes to make these algorithms more energy-efficient by using more efficient implementations of primitive operations. This is an interesting position paper on integrated algorithm design and hardware design. The second article, by Lissandrini et al., argues that exploitation of information graphs can lead to novel query answering capabilities that go beyond the existing capabilities of keyword search. In particular, it focuses on *exemplar queries* on information graphs, a new querying paradigm where a user keyword query on an information graph is treated as an example from a desired result set and the system infers other elements of the result set via similarity search and ranking.

We are pleased to resume the Distinguished Profiles column with two featured articles in this issue. The first article features the interview with Kian-Lee Tan, who is a Provost's Chair Professor of Computer Science at the National University of Singapore (NUS) and the Vice Dean for Research in the School of Computing. In this interview, Kian-Lee shares with us his past and recent research projects, his view on how to build a strong database group at NUS (without pushing students too hard), and time management tips that allowed him to balance work, family, and church activities. The second interview features Sudipto Das, who graduated from the University of California Santa Barbara and won the 2013 SIGMOD Jim Gray Doctoral Dissertation Award. Sudipto talks about his dissertation research on "Scalable and Elastic Transactional Data Stores for Cloud Computing Platforms," as well as how to deal with ups and downs in PhD studies.

In the Research Centers Column, Bentayeb et al. outline the research agenda at the ERIC laboratory in Lyon, France, which includes two research teams: decision support information systems (SID) and data mining and decision (DMD). The SID team addresses a range of applications and research issues in data warehousing and OLAP, including the integration of complex data sources into an Active XML repository, exploiting active rules and mining logged events to automate integration tasks, supporting OLAP on textual data, and developing a cloud computing environment for big data warehousing and OLAP. The DMD team aims to create new systems, models, and algorithms for data mining and decision making, including efforts on topic modeling for dealing with textual datasets extracted from the Web and leveraging topological graphs to design metrics well-suited for machine learning.

In the Industry Perspectives Column, Ellis and his colleagues at IBM Research present the viewpoint that much work has been done to process big data, but more needs to be done to *understand* big data. As an

effort to bridge the gap, they present Helix, a system for guided exploration of big data. Helix provides a unified view of sources, ranging from spreadsheets and XML files with no schema, to RDF graphs and relational data with well-defined schemas. Helix users explore these heterogeneous data sources through a combination of keyword searches and navigation of linked web pages. This article also presents a set of real-world usage scenarios and the lessons learned in the course of developing Helix.

This issue features an event report, by Pedersen, Castellanos, and Dayal, on the Seventh International Workshop on Business Intelligence for the Real time Enterprise (BIRTE 2013) co-located with the VLDB 2013 conference. As business analytics evolves to address new challenges such as big data with high velocity and predictive analytics, BIRTE is becoming an important venue for researchers to present recent and ongoing work in this domain. BIRTE 2013 featured an exciting technical program including two keynotes, an invited industrial talk, a panel, and a number of peer-reviewed papers from different countries in Europe, Africa, and Asia.

This issue closes with the call for papers for the Special Issue on Visionary Ideas in Data Management, to be published as the June 2015 Issue of the ACM SIGMOD Record.

On behalf of the SIGMOD Record Editorial board, I hope that you all enjoy reading the December 2014 issue of the SIGMOD Record and wish you a happy, productive year in 2015!

Your submissions to the Record are welcome via the submission site:
http://sigmod.hosting.acm.org/record

Prior to submitting, you are encouraged to read the Editorial Policy on the SIGMOD Record's Web site (http://www.sigmod.org/publications/sigmod-record/sigmod-record-editorial-policy).

Yanlei Diao

December 2014

## Past SIGMOD Record Editors:

Harrison R. Morse (1969)
Daniel O'Connell (1971 – 1973)
Randall Rustin (1974-1975)
Douglas S. Kerr (1976-1978)
Thomas J. Cook (1981 – 1983)
Jon D. Clark (1984 – 1985)
Margaret H. Dunham (1986 – 1988)
Arie Segev (1989 – 1995)
Jennifer Widom (1995 – 1996)
Michael Franklin (1996 – 2000)
Ling Liu (2000 – 2004)
Mario Nascimento (2005 – 2007)
Alexandros Labrinidis (2007 – 2009)
Ioana Manolescu (2009-2013)