

Editor's Notes

Welcome to the December 2021 issue of the ACM SIGMOD Record!

This issue starts with the Database Principles column featuring an article by Cormode on current trends in data summaries. The focus of data summarization is on finding small data structures, such as samples or sketches, that compactly represent large data sets, can be updated flexibly, and can answer accurately certain queries on the original data. Key application areas include approximate query processing, as well as distributed and stream processing. The article draws an *approximate summary* of efforts in this area, surveying the topics of summaries for machine learning, machine learning for summaries, summaries in privacy, robust streaming, and approximate counting. The author also discusses new bounds and new applications for existing summaries, and outlines open problems for future research.

The Vision column features two articles. The first article, by Mansour, Srinivas, and Hose, aims to address major open problems of handling artifacts in Open Data Science, focusing on which artifacts should be combined to achieve user goals and on how to find artifacts that are semantically similar or connected. The authors propose a federated data-science platform called KEK that closes the gap, by enabling automatic artifact location and sharing, thus breaking down silos in data science. The article details the platform and lists open research challenges and opportunities in this space.

The second article in the Vision column, by Amer-Yahia and colleagues, introduces an end-to-end data-exploration system that is able to guide users in the exploration process, by being reactive and anticipative both for data discovery and data linking. Systems with such capabilities have the potential to open productive data exploration to users with different domain and data-science expertise and experiences in various scientific communities. The system described in the article, called INODE, leverages both machine learning and semantics for data management, encapsulating domain semantics in exploration by example, by natural language, and by recommendation. The authors describe the INODE architecture and discuss challenges and opportunities that arise from the project.

The Research Articles column features an article by Bonifati, Mior, Naumann, and Noack that provides an analysis of participation of women in papers at various top-level database conferences and journals. The preliminary findings of the authors show that there is an overall growth of the number of accepted papers authored by women in major database conferences. The study also examines how the data-management field stands with respect to the fields of HCI, AI, Algorithms, Networking, and Operating Systems. The entire analysis presented in the article is reproducible, with the code and additional results being publicly available.

The Advice to Mid-Career Researchers column presents an article by Balazinska, who invites the readers promoted to senior roles to pause and reflect on where they are and what the next steps are. The issues discussed in the article include hard work in doing great things, opportunities to grow in research and to expand the types of impact that one can have, growth in mentoring of more junior researchers, and leadership in the community. The article also contains advice on overcoming stress.

The article by Arulraj published in the DBrainstorming column discusses opportunities and challenges in effective and efficient automatic video analytics. The author focuses on difficulties in processing video-analytics queries whose aim is to detect actions, as well as in training filters for each unique combination of parameters in video DBMS. The article outlines ideas for overcoming these

challenges and calls for synergistic solutions by the database, computer-vision, and machine-learning communities.

The Distinguished Profiles column features two articles. The first interview is with Juliana Freire, ACM Fellow and professor at New York University. Juliana has a Google Faculty Research Award, an IBM Faculty Award, and an NSF Career Award. She has served as chair of SIGMOD until earlier this year; her Ph.D. is from Stony Brook. Juliana begins the interview by talking about her experience of battling against outdated traditions in the database-research community. She shares her thoughts on why reproducibility is needed in the fast-changing area of computer science, lists reproducibility barriers, and speaks to the imperatives of keeping research artifacts working and ensuring access to their provenance. Juliana touches on the practical impact of her VisTrails system, talks about the ReproZip project, and outlines her vision of the future of tool-based reproducibility. She then discusses the steps she has taken toward the goal of letting all flowers bloom in the community research, and shares advice for fledgling and mid-career database researchers.

The second interview in the column is with Huanchen Zhang, the 2021 winner of the ACM SIGMOD Jim Gray Dissertation award. After a postdoc at Snowflake, Huanchen is an assistant professor at Tsinghua University; his Ph.D. is from Carnegie Mellon University. In the interview, Huanchen talks about his thesis entitled Memory-Efficient Search Trees for Database Management Systems, including the three-step recipe for designing new search-tree data structures that are compact in size and, at the same time, very fast. He discusses the impact of his thesis work in industry, and provides advice for today's graduate students.

The Reports column features an article by Bonnet, Dong, Naumann, and Tözün about the experience of designing and organizing VLDB 2021 as a hybrid conference. VLDB 2021 took place in August 2021, with 180 in-person attendees in Copenhagen, Denmark and with 840 remote attendees. The article describes the key decisions of the general chairs and PC chairs of the conference, and shares the lessons learned. The authors believe that the hybrid format for scientific conferences is here to stay and opens up new opportunities for everyone.

On behalf of the SIGMOD Record Editorial board, I hope that you enjoy reading the December 2021 issue of the SIGMOD Record!

Your submissions to the SIGMOD Record are welcome via the submission site:

<https://mc.manuscriptcentral.com/sigmodrecord>

Prior to submission, please read the Editorial Policy on the SIGMOD Record's website:

<https://sigmodrecord.org/sigmod-record-editorial-policy/>

Rada Chirkova

December 2021

Past SIGMOD Record Editors:

Yanlei Diao (2014-2019)	Ioana Manolescu (2009-2013)	Alexandros Labrinidis (2007-2009)
Mario Nascimento (2005-2007)	Ling Liu (2000-2004)	Michael Franklin (1996-2000)
Jennifer Widom (1995-1996)	Arie Segev (1989-1995)	Margaret H. Dunham (1986-1988)
Jon D. Clark (1984-1985)	Thomas J. Cook (1981-1983)	Douglas S. Kerr (1976-1978)
Randall Rustin (1974-1975)	Daniel O'Connell (1971-1973)	Harrison R. Morse (1969)