Letter from the Editor-in-Chief

Thank You, David!

I know I represent the readers, the associate editors, and also the broad database community when I say we are extremely grateful to David Lomet for his distinguished and dedicated service as the Editor-in-Chief of the Data Engineering Bulletin for the last 26 years.

Since its launch in 1977, the Bulletin has produced a total of 154 issues. Reading through the topics of the past issues that spanned more than four decades makes me feel nothing short of amazing. They show not just how far the database research has come, but to a certain extent, how much the entire field of computer science and the IT industry have evolved. While important topics never fail to arise in the Bulletin in a timely fashion, it is also interesting to observe in the 154 issues many recurring topics, including query optimization, spatial and temporal data management, data integration, etc. It proves that the database research has a solid foundation that supports many new applications, and at the same time, it demonstrates that the database research is constantly reinventing itself to meet the challenges of the time. What the Bulletin has faithfully documented over the last 42 years is nothing else but this amazing effort.

Among the 154 issues since the launch of the Bulletin, David had been the Editor-in-Chief for 103 of them. This itself is a phenomenal record worth an extra-special celebration. But more importantly, David shaped the discussions and the topics in the long history of the Bulletin. I had the honor to work with David in 2016 and 2017 when I served as the associate editor for two Bulletin issues. What was most appealing to me was the opportunity of working with the top experts on a topic that I am passionate about. The Bulletin is truly unique in this aspect.

I understand the responsibility and the expectation of the Editor-in-Chief, especially after David set such a great example in the last 26 years. I thank David and the associate editors for their trust, and I look forward to working with authors, readers, and the database community on the future issues of the Data Engineering Bulletin.

The Current Issue

Machine learning is changing the world. From time to time, we are amazed at what a few dozen lines of python code can achieve (e.g., using PyTorch, we can create a simple GAN in under 50 lines of code). However, for many real-life machine learning tasks, the challenges lie beyond the dozen lines of code that construct a neural network architecture. For example, hyperparameter tuning is still considered a "dark art," and having a platform that supports parallel tuning is important for training a model effectively and efficiently. Model training is just one component in the life cycle of creating a machine learning solution. Every component, ranging from data preprocessing to inferencing, requires just as much support on the system and infrastructure level.

Joseph Gonzalez put together an exciting issue on the life cycle of machine learning. The papers he selected focus on systems that help manage the process of machine learning or resources used in machine learning. They highlight the importance of building such supporting systems, especially for production machine learning platforms.

Haixun Wang
WeWork Corporation