

An Interview with Dr. Shipeng Yu, Winner of ACM SIGKDD 2021 Service Award

ABSTRACT

Shipeng Yu, Ph.D. is the recipient of the 2021 ACM SIGKDD Service Award, which is the highest service award in the field of knowledge discovery and data mining. Conferred annually on one individual or group in recognition of outstanding professional services and contributions to the field of knowledge discovery and data mining, Dr. Yu was honored for his years of service and many accomplishments as general chair of KDD 2017 and currently as sponsorship director for SIGKDD. Dr. Yu is Director of AI Engineering, Head of the Growth AI team at LinkedIn, the world's largest professional network. He sat down with SIGKDD Explorations to discuss how he first got involved in the KDD conference in 2006, the benefits and drawbacks of virtual conferences, his work at LinkedIn, and KDD's place in the field of machine learning, data science and artificial intelligence.

CONGRATULATIONS ON RECEIVING THE SIGKDD SERVICE AWARD FOR YOUR CONTRIBUTIONS AS GENERAL CHAIR AND SPONSORSHIP DIRECTOR. PLEASE TELL US ABOUT YOURSELF. HOW DID YOU GET INTO THE FIELD OF DATA SCIENCE, AI AND MACHINE LEARNING?

I was born and raised in China. I became interested in math early on, and studied the subject in undergraduate and graduate schools. I pursued my doctoral work at the University of Munich, where I had the good fortune to study under renowned Data Mining and AI experts Professor Hans-Peter Kriegel and Dr. Volker Tresp. While in Germany, I had the opportunity to work at Siemens Corporate Research, exploring broad applications of general machine learning, though much of my work focused on statistical machine learning. Given my experience in machine learning and data mining at a large scale, after completing my Ph.D., I accepted a role at Siemens Healthcare in the U.S. Over a span of eight years in the healthcare sector, I was introduced to many different kinds of medical data: image data, clinical data, genetic data. The goal was to improve diagnoses and personalize treatment through data mining.

AS THE HEAD OF GROWTH AI AT LINKEDIN, CAN YOU SHARE WHAT THAT ROLE LOOKS LIKE?

While I appreciated the very real impact—you are literally saving lives—of working with medical data, regulatory and other challenges are prevalent in the healthcare industry. You are also dealing with small data in most cases. My move to LinkedIn in early 2015 was prompted by a desire to apply the tenets of data mining in a new environment and a new industry. The Growth AI team at LinkedIn has broad-based responsibility, including network and engagement growth, as well as retention of the platform's 800 million members. We are primarily concerned with building the

right network and getting the right communications to the right members at the right time. Data mining and machine learning are essential tools to optimize these growth charters. I still do research and publish papers, but that work is also product focused, with a goal to improve business functions and strategies.

AS SOMEONE WITH EXPERIENCE IN THE PRIVATE SECTOR AND ACADEMIC RESEARCH, WHERE CAN ACADEMIA AND INDUSTRY HELP ONE ANOTHER?

With experience spanning both academia and industry, I can say there are definitely some not-so-subtle differences between the two. Without the rigid pressure of metrics or budgets, universities tend to focus on fundamental research—things that are longer-term, more groundbreaking. Access to government funding and a pool of student researchers also contributes to the success and nature of academic research. On the other hand, industry has customers, partners, shareholders and other stakeholders to satisfy. As a result, businesses tend to focus research on solving fascinating real-world problems, often with a more aggressive timeline. While perhaps different in approach or context, I do believe there are problems common to both academia and industry, that may be tackled with data science. Because KDD spans both worlds, the organization and its many volunteers are ideally suited to help improve collaboration between higher education and the business sector, making the field of data science more equitable and diverse.

WHEN DID YOU FIRST DISCOVER KDD AND HOW HAS YOUR RELATIONSHIP WITH THE CONFERENCE EVOLVED?

KDD came on my radar in my early years at graduate school when some of my work was published at KDD conferences. I really got my foot in the door when I volunteered to be a reviewer at the first conference I attended, KDD 2006. That's when I began to appreciate the international community of data mining and knowledge discovery researchers, and offered to help facilitate conference workshops. Between 2010 and 2016, my involvement grew steadily, and I was thrilled to be part of planning the first KDD conference in a major U.S. city, KDD 2016 in San Francisco. A resounding success, that year there were more than 2,700 attendees from 88 countries.

I then had the honor and privilege to serve as general chair for KDD 2017, which took place in Halifax, Nova Scotia. A fairly remote location and last-minute venue change posed challenges from an organizing perspective. However, building on the success of the prior year and adopting many of the same tactics to recruit volunteers and attract registrants, we were pleased to draw more than 1,700 attendees. Since serving as the sponsorship chair for the 2018 conference in London, I have taken on responsibility for leading SIGKDD's sponsorship efforts.

AS SPONSORSHIP DIRECTOR FOR SIGKDD, WHAT ARE YOUR OBJECTIVES AND HOW DO YOU MEET THEM?

Sponsorships have been key to the longevity and success of the conference, with generous financial support ensuring KDD remains home to some of the most highly cited research papers in data science. While previous sponsorship chairs served one-year stints, I believe continued growth in this area requires leadership continuity, so have been pleased to serve in this role for three years. With the organization surpassing \$1 million in annual sponsorship—\$1.2 million in 2018 and \$1.1 million in 2019, my goal is to continue and accelerate this trajectory.

I am proud that we have secured interest from both government and corporate sponsors over my tenure. This healthy mix of support has enabled us to set sponsorship records. The switch from a physical package and to a virtual package necessitated by the Coronavirus pandemic, coupled with resulting economic uncertainty, have posed some sponsorship struggles. Fortunately, though, we have been building long-term relationships with many of our sponsors, and, as a result, financial support for our virtual conference improved in 2021.

YOU HAVE HELPED ORGANIZE MULTIPLE IN-PERSON CONFERENCES. HOW HAVE THE PANDEMIC AND THE SHIFT TO VIRTUAL CONFERENCES CHANGED THE EXPERIENCE FOR ATTENDEES?

The coronavirus pandemic has affected us in ways we never imagined. Every organization has had to adapt to the new normal. Prior to the public health crisis, KDD thrived as a vibrant, physical, in-person event. The conference was at its best when people were talking and collaborating in small group discussions. We have been forced to tackle head on the uncertainty around how we can recreate

that experience online, how sponsors can reach members in this format, and which offerings will and won't work virtually. At the same time, virtual conferences have opened up a lot of opportunities for attendees to interact with speakers and authors. It also democratizes the conference, making it available for everyone, as registration fees have been reduced and travel isn't required. That's why I believe there is value in the hybrid format. Yes, it had been done in the past, but it wasn't mainstream. Now people can see that a combination of virtual and in-person conference setup not only works but expands opportunities for people to get together in a way that makes sense for them.

WHY DO YOU THINK IT IS IMPORTANT TO VOLUNTEER FOR SIGKDD?

KDD is special not only because of the groundbreaking discoveries presented at the show, but also the connections and bonds we are able to form with our colleagues. I have been extremely fortunate to serve as general chair as well as my current role. Being the sponsorship director at KDD is not just about raising money. It's about raising awareness around our field and the phenomenal international community that we have built and exponentially grown over the past few years. Volunteering creates new and expanded opportunities to sustain and strengthen valuable personal and professional relationships, and give back to an organization at the forefront of advancing our industry.

LASTLY, WHAT DOES THE SERVICE AWARD WIN MEAN TO YOU?

I cannot begin to express my gratitude to everyone who has helped contribute to KDD's growth and success over the past two decades. I am honored to join such an accomplished list of professionals and servants who have won this award, including last year's winner, Dr. Michael Zeller. I encourage everyone in the data mining and knowledge discovery community to join us next year at KDD 2022.