

Letter from the Editor-in-Chief

In this issue, we continue our discussions on data science. Following Jeff Ullman’s opinion piece in the June 2020 issue titled “The battle for data science,” we feature an opinion piece from M. Tamer Özsu. Özsu brings up a few critical questions about data science ranging from “Is data science a discipline?”, “Who is a data scientist?” to “Who owns data science?” These questions are important not only for academia but also in the tech industry, where “data science” remains a vague term to this day. Different tech companies often define data science in different ways, resulting in an identity crisis for data scientists as they balance their roles as applied scientists, machine learning engineers, statisticians, data analysts, data engineers, etc. More importantly, as the discipline is not well-defined, tools, platforms, and eco-systems cannot be “standardized,” which means they are unlikely to be perfected. Please read Özsu’s opinion piece “A Systematic View of Data Science.”

Lei Chen organized this special issue on the topic of human-powered AI or human-in-the-loop AI. While these are not exactly new topics as there has been a lot of work in areas ranging from crowdsourcing to active learning, the topic of making AI and humans working together and benefiting each other is gaining a lot of traction. In this issue, we look into problems such as how to divide labor between humans and machines, how to utilize crowdsourcing on federate learning, how to provide end-to-end solution in human-in-the-loop systems, how to use AI to model human behavior, and how human-powered AI can help with online misinformation detection.

Haixun Wang
Instacart