C. F. JEFF WU

Georgia Institute of Technology

https://en.wikipedia.org/wiki/C._F._Jeff_Wu

Data science in the age of uncertainties

Biography

F. Jeff Wu is Professor and Coca Cola Chair in Engineering Statistics at the School of Industrial and Systems Engineering, Georgia Institute of Technology. He was elected a Member of the National Academy of Engineering (2004), and a Member (Academician) of Academia Sinica (2000). A Fellow of American Society for Quality, of Institute of Mathematical Statistics, of INFORMS, and of American Statistical Association. He received the COPSS Presidents' Award in 1987, COPSS Fisher Lecture Award in 2011, Deming Lecture in 2012. He has won other awards, including the Shewhart Medal (2008), the Pan Wenyuan Technology Award (2008), Class of 1934 Distinguished Professor Award and Sigma Xi Monie A. Ferst Award both at Georgia Institute of Technology in 2020. He was the 1998 Mahalanobis Memorial Lecturer at the Indian Statistical Institutes, received the inaugural Akaike Memorial Lecture Award in 2016 sponsored by the Japan Statistical Society and the Institute of Statistical Mathematics, Tokyo, the 2017 Box Medal from ENBIS, and an honorary doctor degree at the University of Waterloo.

He has published more than 185 research articles. He has supervised 50 Ph.D.'s, out of which more than half are teaching in major research departments in statistics/engineering/business in US/Canada/Asia/Europe. Among them, there are 22 Fellows of ASA, IMS, ASQ, IAQ, and IIE. He co-authors with Mike Hamada the book "Experiments: Planning, Analysis, and Optimization" (Wiley, 3rd ed. 2021) and with R. Mukerjee the book "A Modern Theory of Factorial Designs" (Springer, 2006).

Abstract

In this talk I will first trace the history of my role in the coining of the term "data science". For many years since the early 1980s, I had grown dissatisfied with using the term "statistics" to describe my profession because it is usually connected with descriptive statistics, while what statisticians do can be summarized as a *trilogy* of data collection, data modeling, and problem solving. Thus I proposed the terms data science and data scientist in a public lecture at the U. of Michigan in 1997. With the explosion of huge data collected through the internet, data science has grown to become a very popular, fashionable and impactful profession. I will describe how it is so different from the traditional meaning and work of statistics. A major new component is the role played by computer scientists and the new emphases on algorithms, coding and huge data they bring in. I will end with some examples about the applications to uncertainty quantification and variation reduction.