



Association for
Computing Machinery

Advancing Computing as a Science & Profession



NEWS RELEASE

CONTACT: Jim Ormond
212-626-0505
ormond@hq.acm.org

LEADING EXPERTS EXAMINE HOW DATA SCIENCE AND AI ARE RESHAPING TECHNOLOGY AND TRANSFORMING SOCIETY

ACM-IMS Gathering on the Foundations of Data Science Explores Deep Learning, Ethics, Fairness and the Future of AI

New York, NY, June 6, 2019 – Computing and statistics underpin the rapid emergence of data science as a pivotal multi-disciplinary field today. The two key professional organizations in these areas, the Association for Computing Machinery (ACM) and the Institute of Mathematical Statistics (IMS), have organized the [ACM-IMS Interdisciplinary Summit on the Foundations of Data Science](#) to be held on June 15, 2019 at the Palace Hotel in San Francisco. This one-day event will bring together distinguished speakers and panelists to address topics such as deep learning, reinforcement learning, fairness, ethics, and the future of data science.

Keynote Talks

“Deep Learning for Tackling Real-World Problems”

Jeffrey Dean, Google Senior Fellow and SVP, Research and Health

In this talk, Dean will look at how recent advances based on deep learning have made significant strides in fundamental areas such as computer vision, speech recognition, language understanding and translation. Given these advances, he will then explore how deep learning can now help us tackle some of the major challenges in the world, such as improving access to healthcare and creating new tools for scientific discovery. Finally, he will touch on how deep learning is changing the way in which we think about building computational hardware—creating a resurgence in new and interesting computer architecture work to design systems that target and dramatically accelerate deep learning workloads.

“Machine Learning: A New Approach to Drug Discovery”

Daphne Koller, CEO and Founder, insitro

Modern medicine has given us effective tools to treat some of the most significant and burdensome diseases. At the same time, it is becoming consistently more challenging to develop new therapeutics. A key contributor to this trend is that the drug development process involves multiple steps, each of which involves a complex and protracted experiment that often fails. For many of these phases, we can

develop machine learning models to help predict the outcome of these experiments. The key will be to train powerful ML techniques on sufficient amounts of high-quality, relevant data. In this talk, Koller will discuss how she and her colleagues are developing a new approach to drug development that uses high-quality data and ML models to design novel, safe, and effective therapies that help more people, faster, and at a lower cost.

"Making the Black Box Effective: What Statistics Can Offer?"

Emmanuel Candès, Professor of Mathematics and Statistics at Stanford University

The ongoing data science revolution has been driven by impressive technological advances in the capture, storage, and processing of data, across a wide range of domains. Of particular interest is the recent progress in machine learning which provides us with many potentially effective tools to learn from datasets of ever increasing sizes and make useful predictions. While many scientists and engineers are, for good reasons, slowly getting comfortable with the idea of using models that are extremely difficult to interpret — black boxes if you will — two things cannot be compromised upon. The first is the reproducibility of scientific results. The second concerns the validity of predictions. In this talk, Candès will review recent progress in statistics which addresses these concerns. He will present broad methodologies that can be wrapped around any black box as to produce results that can be trusted.

Panel Discussions

Deep Learning, Reinforcement Learning, and Role of Methods in Data Science

Moderator: *Joseph Gonzalez, University of California, Berkeley*

Panelists: *Shirley Ho, Flatiron Institute; Sham Kakade, University of Washington; Suchi Saria, Johns Hopkins University; and Manuela Veloso, J.P. Morgan AI Research and Carnegie Mellon University*

In addition to rapid advances in raw compute power and our ability to capture data, methodological advances fuel the extraordinary rise of data science. Deep learning and reinforcement learning, machine learning techniques with roots that go back at least 50 years, are currently attracting intense focus. The panel will discuss these particular methods as well as the role of methods and method development in data science more generally.

Robustness and Stability in Data Science

Moderator: *Ryan Tibshirani, Carnegie Mellon University*

Panelists: *Aleksander Madry, MIT; Xiao-Li Meng, Harvard University; Richard J. Samworth, University of Cambridge and The Alan Turing Institute; and Bin Yu, University of California, Berkeley*

The concept of statistical "robustness" has attracted much attention from the statistical research community, dating back to the seminal work of Huber in the 1970's. The key idea is that statistical inferences and predictions should be insensitive to outliers and to small perturbations of the data. Data scientists have recently formulated more general concepts of "stability" that build on the earlier ideas but now in the context of large-scale data analyses. The panel will discuss recent progress and future directions.

Fairness and Ethics in Data Science

Moderator: *Yannis Ioannidis, National and Kapodistrian University Athens*

Panelists: *Joaquin Quiñero Candela, Facebook; Alexandra Chouldechova, Carnegie Mellon University; Andrew Gelman, Columbia University; and Kristian Lum, Human Rights Data Analysis Group*

As data science plays an increasingly important role on our lives, critically important questions of fairness and ethics have taken on a new urgency. One worry is that data-based algorithms might reinforce previous patterns of bias and discrimination. Another concerns the motivations and biases that can permeate the work of (human) data scientists. Yet another is fake news and "deep fakes." The panel will discuss these issues, their ramifications, and possible safeguards.

Future of Data Science

Moderator: *David Madigan, Columbia University*

Panelists: *Michael I. Jordan, University of California, Berkeley; and Jeannette Wing, Columbia University*

The panelists will discuss the most exciting research avenues in data science and what new innovations might be on the horizon.

Steering Committee: Jeannette Wing (Event Co-Chair), David Madigan (Event Co-Chair), Magdalena Balazinska, Joseph Gonzalez, Chris Holmes, Yannis Ioannidis, Ryan Tibshirani, and Daniela Witten

Media Registration: Contact Jim Ormond at ormond@hq.acm.org

About ACM

[ACM, the Association for Computing Machinery](#) is the world's largest educational and scientific computing society, uniting computing educators, researchers and professionals to inspire dialogue, share resources and address the field's challenges. ACM strengthens the computing profession's collective voice through strong leadership, promotion of the highest standards, and recognition of technical excellence. ACM supports the professional growth of its members by providing opportunities for life-long learning, career development, and professional networking.

About IMS

[IMS, the Institute of Mathematical Statistics](#), is an international professional and scholarly society devoted to the development, dissemination, and application of statistics and probability. The Institute currently has about 4,000 members in all parts of the world.

###