



Association for
Computing Machinery

Advancing Computing as a Science & Profession

NEWS RELEASE

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

NEW CONFERENCE SERIES TO BRING TOGETHER TOP COMPUTER SCIENTISTS AND STATISTICIANS TO ADDRESS DATA SCIENCE CHALLENGES

First ACM-IMS Foundations of Data Science Conference Opens Virtually October 19

New York, NY, October 7, 2020 –The Association for Computing Machinery (ACM) together with the Institute of Mathematical Statistics (IMS) will hold the first-ever [ACM-IMS Foundations of Data Science \(FODS\) Conference](#) virtually on October 19-20. This interdisciplinary event will bring together researchers and practitioners to address foundational data science challenges in prediction, inference, fairness, ethics and the future of data science.

“Data science is a new, emerging field, building its foundations from computer science, statistics, and many other quantitative disciplines,” said FODS General Co-chair Jeannette Wing, Columbia University, and Fellow, Association for Computing Machinery. “Big data is not new: through large, one-of-a-kind, expensive instruments, scientists have been collecting and generating massive amounts of data for decades. What *has* changed is that the internet has become an instrument for anyone, not just scientists, to collect and generate data, and that that data is about people. We also have powerful AI, machine learning, and statistical techniques that allow us to *interpret* and *gain value* from the data in new ways. And because so much data is about people, we must address up front questions of ethics and privacy. We are witnessing a new era where every sector, including healthcare and finance, is being transformed by data science. We believe that our interdisciplinary approach to organizing this conference will make it an important research gathering for many years to come.”

“FODS is a first-of its-kind conference in that it is a collaboration between the two leading scientific societies in computing and statistics,” added FODS General Co-chair, David Madigan, Northeastern University, and Fellow, Institute for Mathematical Statistics. “We believe this cross-collaboration between computer scientists and statisticians is the most effective way to foster groundbreaking new research in this field. Building on the success of the initial summit ACM and IMS co-organized in 2019, we have put together an exciting program featuring the world’s top researchers and practitioners. We also hope that the virtual nature of this year’s conference will encourage participants from around the world to engage with us.”

ACM-IMS FODS 2020 HIGHLIGHTS

Keynote Speakers

“AutoML and Interpretability: Powering the Machine Learning Revolution in Healthcare”

Michaela van der Schaar, The Alan Turing Institute

AutoML and interpretability are both fundamental to the successful uptake of machine learning by non-expert end users. This keynote presents state-of-the-art AutoML and interpretability methods for healthcare developed in van der Schaar’s lab and how they have been applied in various clinical settings (including cancer, cardiovascular disease, cystic fibrosis, and recently Covid-19), and then explains how these approaches form part of a broader vision for the future of machine learning in healthcare.

“Semantic Scholar, NLP, and the Fight Against COVID-19”

Oren Etzioni, Allen Institute for AI (AI2)

Etzioni’s talk will describe the dramatic creation of the COVID-19 Open Research Dataset (CORD-19) at the Allen Institute for AI and the broad range of efforts, both inside and outside of the Semantic Scholar project, to garner insights into COVID-19 and its treatment based on this data. The talk will highlight the difficult problems facing the emerging field of Scientific Language Processing.

FODS 2020 Papers (Partial List)

For a list of all accepted papers, [visit here](#).

“Incentives Needed for Low-Cost Fair Data Reuse”

Roland Maio, Augustin Chaintreau, Columbia University

One of the central goals in algorithmic fairness is to build systems with fairness properties that compose gracefully. Although the importance of this goal was recognized early, limited progress has been made. In this work, Maio and *Chaintreau* propose a fresh approach to building fairly composable data-science pipelines by incorporating information about parties' incentives into fairness interventions. Their results open several new directions for research on fair data-science pipelines, fair machine learning, and algorithmic fairness more broadly.

“Applying Algorithmic Accountability Principles and Frameworks to Ecosystem Forecasting: A Case Study in Forecasting Shellfish Toxicity in the Gulf of Maine”

Isabella Grasso, David Russell, Jeanna Matthews, Clarkson University; Abigail Matthews, University of Wisconsin-Madison; Nick Record, Bigelow Laboratory for Ocean Sciences

Ecological forecasts are used to drive decisions that can have significant impacts on the lives of individuals and on the health of ecosystems. In this paper, the authors discuss their experience with applying algorithmic accountability principles and frameworks to ecosystem forecasting, in particular to forecasting shellfish toxicity in the Gulf of Maine using a dataset produced by the Marine Biotxin Monitoring Program conducted by the Department of Marine Resources (DMR).

“StyleCAPTCHA: CAPTCHA based on style-transferred images to defend against Deep Convolutional Networks”

Haitian Chen, Bai Jiang and Hao Chen

CAPTCHA has found widespread applications for bot detection in the cyberspace. Many CAPTCHAs are based on visual perception tasks such as text recognition, objection recognition and image classification. However, they are under serious threat from modern visual perception technologies, especially deep convolutional networks (DCNs). The authors propose a novel CAPTCHA, called StyleCAPTCHA, which asks users to classify stylized human versus animal face images. Each stylized image in StyleCAPTCHA is created by combining the content representations of a human or animal face image and the style representations of a style reference image, both of which are hidden from the user.

Tutorials

“Causal Reasoning Tutorial”

David Blei, Columbia University

Blei is a professor of Statistics and Computer Science at Columbia University. He is also a member of the Columbia Data Science Institute. He works in the fields of machine learning and Bayesian statistics.

“Fairness, Privacy and Ethics in Data Science Tutorial”

Michael Kearns, University of Pennsylvania

Kearns is a professor of Computer and Information Science at University of Pennsylvania. He is also the Founding Director of the Warren Center for Network and Data Sciences at the University of Pennsylvania. His research interests include topics in machine learning, algorithmic game theory and microeconomics, computational social science, and quantitative finance and algorithmic trading.

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 the leading organization fostering the development and dissemination of the theory and applications of statistics.

###