



Institute for Data Sciences and Engineering
COLUMBIA UNIVERSITY

SEMINAR EVENT

DATE: TUESDAY, OCT. 15

TIME: 3:00 P.M.

LOCATION:
DAVIS AUDITORIUM,
ROOM 412, SHAPIRO
CEPSR



SPEAKER: **Michael I. Jordan**
University of California, Berkeley

TITLE: **On the Computational and Statistical Interface and “Big Data”**

DATE: TUESDAY, OCT. 15

TIME: 3:00 P.M.

LOCATION:
DAVIS AUDITORIUM,
ROOM 412, SHAPIRO
CEPSR

Abstract:

The rapid growth in the size and scope of datasets in science and technology has created a need for novel foundational perspectives on data analysis that blend the statistical and computational sciences. That classical perspectives from these fields are not adequate to address emerging problems in "Big Data" is apparent from their sharply divergent nature at an elementary level---in computer science, the growth of the number of data points is a source of "complexity" that must be tamed via algorithms or hardware, whereas in statistics, the growth of the number of data points is a source of "simplicity" in that inferences are generally stronger and asymptotic results can be invoked.

I present three research vignettes on topics at the computation/statistics interface, the first involving the deployment of resampling methods such as the bootstrap on parallel and distributed computing platforms, the second involving large-scale matrix completion, and the third introducing a methodology of "algorithmic weakening," whereby hierarchies of convex relaxations are used to control statistical risk as data accrue. [Joint work with Venkat Chandrasekaran, Ariel Kleiner, Lester Mackey, Purna Sarkar, and Ameet Talwalkar].

500 WEST 120TH STREET
NEW YORK, NY 10027

PHONE: 212-854-5660
IDSE.COLUMBIA.EDU