



THE POWER OF **BIG DATA**

**A SERIES OF EVENTS PRESENTED BY
TEMPLE UNIVERSITY'S BIG DATA INSTITUTE**

FRIDAY, FEBRUARY 19, 2016

10:00 TO 11:30AM

ROOM 306, SCIENCE EDUCATION RESEARCH CENTER

TEMPLE UNIVERSITY

THE PREDICTIVE POWER OF MASSIVE DATA ABOUT OUR FINE-GRAINED BEHAVIOR

FOSTER PROVOST

**ANDRE MEYER FACULTY FELLOW, AND PROFESSOR OF DATA SCIENCE AND INFORMATION SYSTEMS
NEW YORK UNIVERSITY**

What really is it about “big data” that makes it different from traditional data? In this talk I illustrate one important aspect: massive ultra-fine-grained data on individuals' behaviors holds remarkable predictive power. I examine several applications to marketing-related tasks, showing how machine learning methods can extract the predictive power and how the value of the data "asset" seems different from the value of traditional data used for predictive modeling.

I then dig deeper into explaining the predictions made from massive numbers of fine-grained behaviors by applying a counter-factual framework for explaining model behavior based on treating the individual behaviors as evidence that is combined by the model. This analysis shows that the fine-grained behavior data incorporate various sorts of information that we traditionally have sought to capture by other means. For example, for marketing modeling the behavior data effectively incorporate demographics, psychographics, category interest, and purchase intent.

Finally, I discuss the flip side of the coin: the remarkable predictive power based on fine-grained information on individuals raises new privacy concerns. In particular, I discuss privacy concerns based on inferences drawn about us (in contrast to privacy concerns stemming from violations to data confidentiality). The evidence-counterfactual approach used to explain the predictions also can be used to provide online consumers with transparency into the reasons why inferences are drawn about them. In addition, it offers the possibility to design novel solutions such as a privacy-friendly "cloaking device" to inhibit inferences from being drawn based on particular behaviors.

FOX.TEMPLE.EDU/BIGDATA

