Big Data and HR:
3 Lessons Learned

Prasanna (Sonny) Tambe
FRB-Atlanta, July 20, 2016
“people analytics will ultimately have a vastly larger impact on the economy than the algorithms that now trade on Wall Street or figure out which ads to show us.”

- E. Brynjolfsson in The Atlantic
“Digital exhaust” is creating a revolution in workforce science

Real time labor market information
Mobile phone/GPS/Location data
Web links/Blog references/Facebook
Socio-metric badges
Email network data
Employee referrals
Internal digital chatter
MOOC assessments
Behavioral games
Internal knowledge boards
Discussion board posts
Open source contributions
Online databases of resumes
Trace data from wearable devices
Data vs. Intuition

HURRICANE FRANCES was on its way, barreling across the Caribbean, threatening a direct hit on Florida's Atlantic coast. Residents made for higher ground, but far away, in Bentonville, Ark., executives at Wal-Mart Stores decided that the situation offered a great opportunity for one of their newest data-driven weapons, something that the company calls predictive technology.

A week ahead of the storm's landfall, Linda M. Dillman, Wal-Mart's chief information officer, pressed her staff to come up with forecasts based on what had happened when Hurricane Charley struck several weeks earlier. Backed by the trillions of bytes' worth of shopper history that is stored in Wal-Mart's data warehouse, she felt that the company could "start predicting what's going to happen, instead of waiting for it to happen," as she put it.
The current frontier: Real-time labor supply

Source: Tambe 2014, LinkedIn
The current frontier: Real-time labor demand
Movement from weak to strong signals of individual job performance.
Strong signals of on-the-job performance
What can **online activities** tell us about workers?
CODE ON THE ROAD

Want to put your coding skills to the test? Flex your hacker skills for a chance to chat with the Uber team!
Question 2

You are helping design our dispatch system. When a trip is requested, you need to return the driver with the shortest ETA. If a driver canceled, the next driver with the shortest ETA out of k ETAs is dispatched instead.

Which data structure would you use to store the k drivers and dispatch the driver with the shortest ETA?

[A] Array

[B] Heap

[C] Hash Table

[D] Binary Search Tree
What about “EQ” skills?
Better management of information work
Better measures of online and offline communication
Access to information diffusion predicts individual productivity.

- Each additional ‘keyword seen’ is associated with about $70 of additional revenue generated.

Seeing information sooner also predicts higher productivity.

- An additional word seen within the first week of its emergence in the network is worth ~ $321.
- An additional word seen within the first month of its emergence in the network is worth ~ $115.

Source: * Aral, Brynjolfsson & Van Alstyne “Productivity Effects of Information Diffusion in Networks.”
New battles over worker privacy.
We already trade *privacy for discounts* in many consumer markets
Like with credit histories, *Opting-out* may not be a choice.
The light and dark sides of “big data and HR”