

# Soc 596: Computational Social Science

Matthew J. Salganik  
Department of Sociology  
Princeton University

01-01-Introduction to computational social science



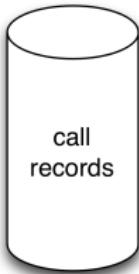
# What is computational social science?

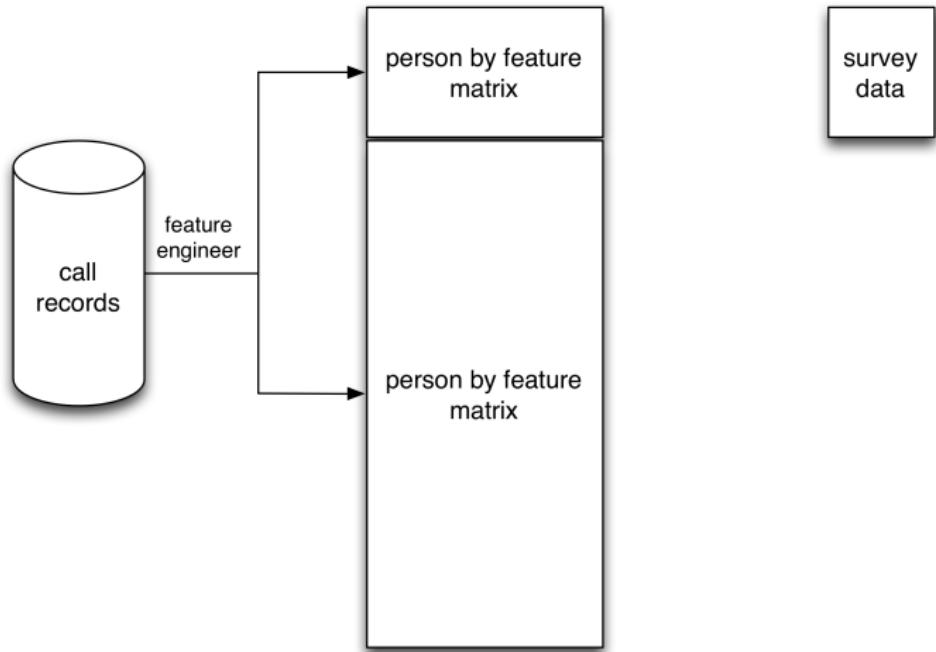
# Predicting poverty and wealth from mobile phone metadata

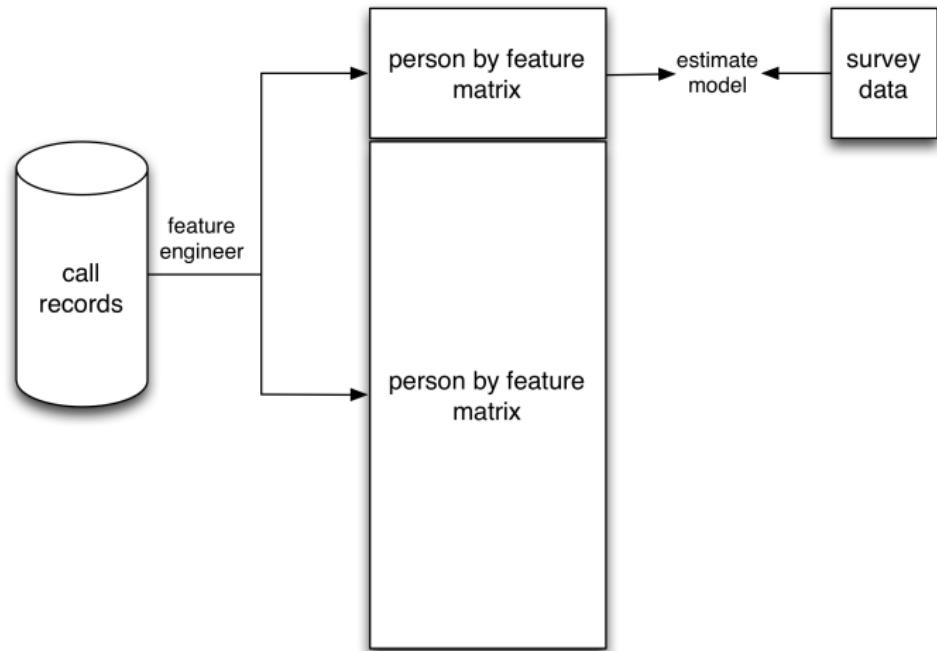
Joshua Blumenstock,<sup>1\*</sup> Gabriel Cadamuro,<sup>2</sup> Robert On<sup>3</sup>

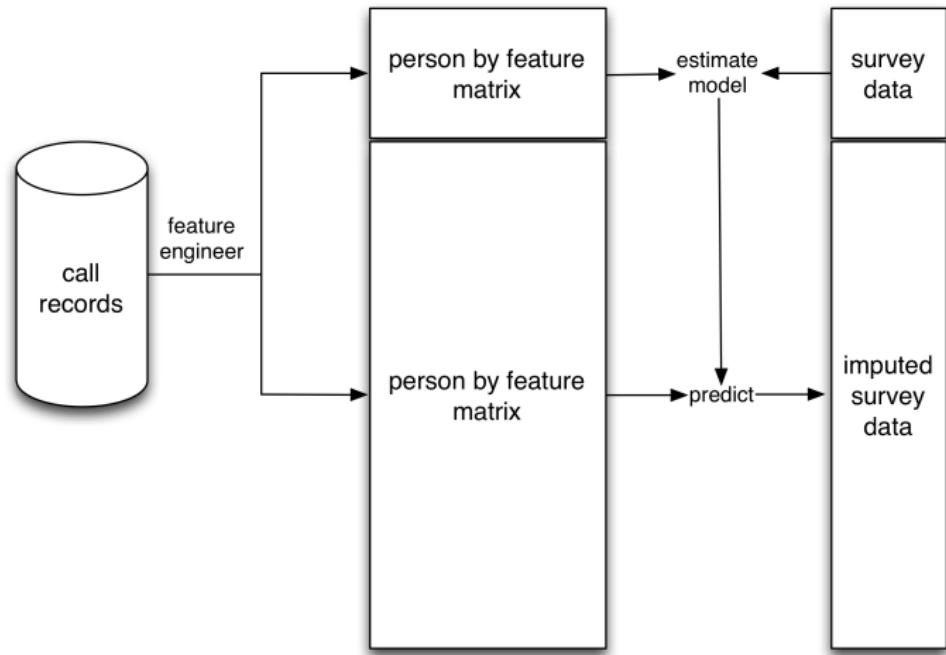
<http://dx.doi.org/10.1126/science.aac4420>

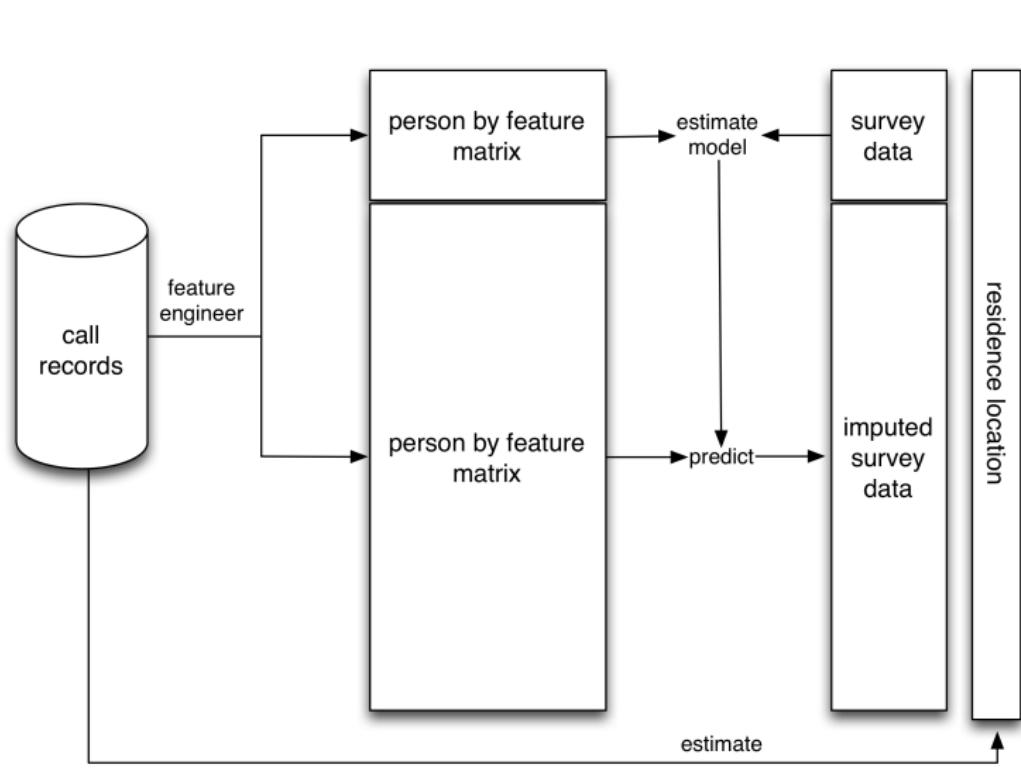


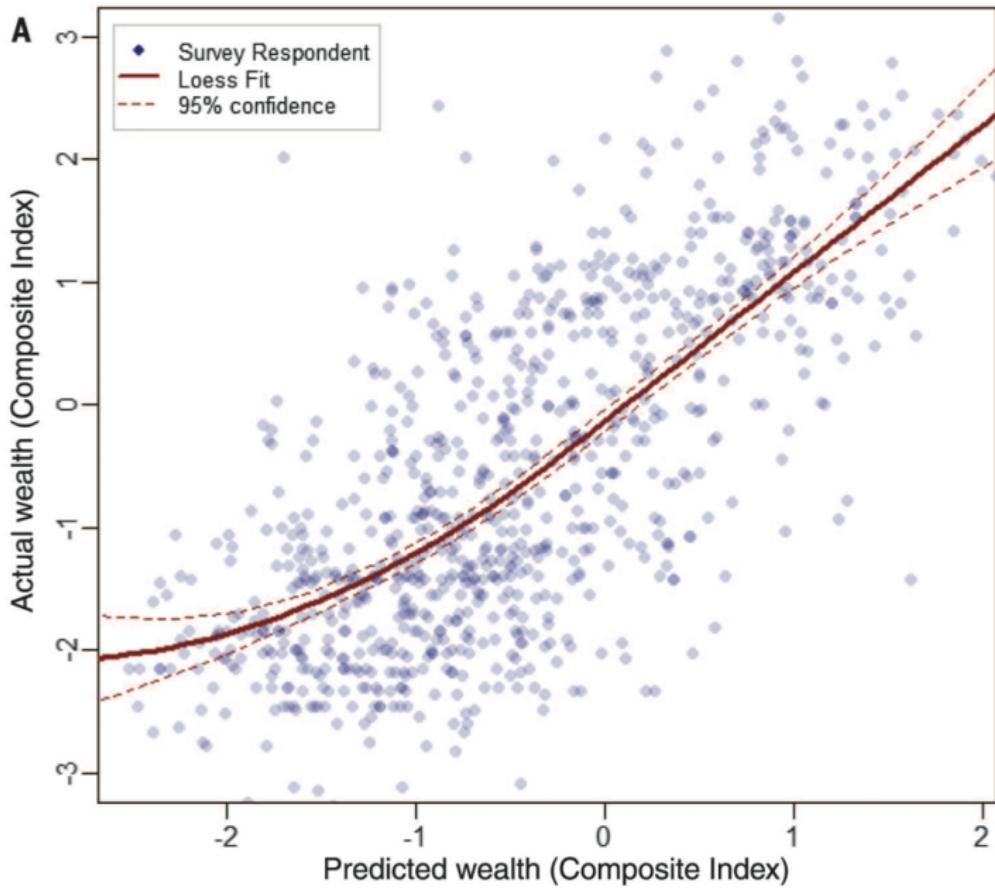


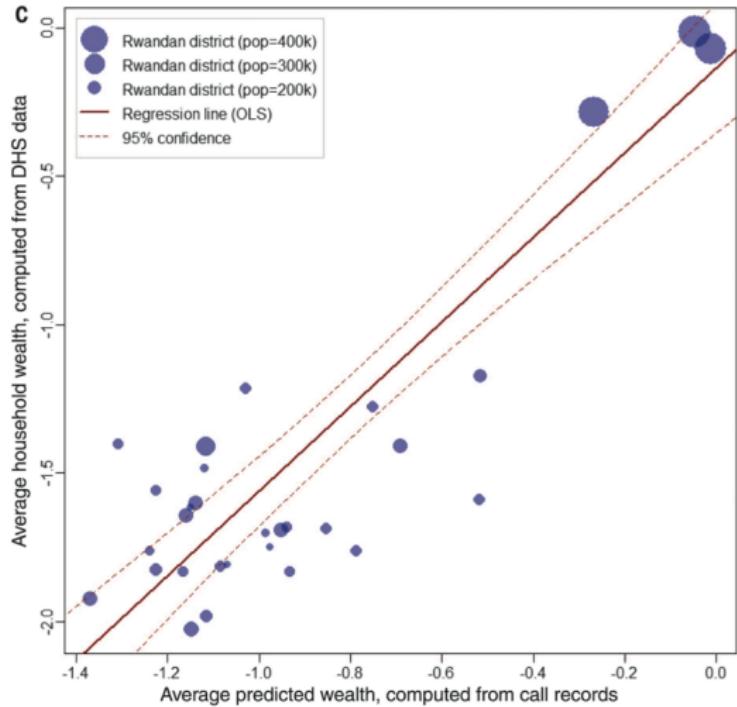


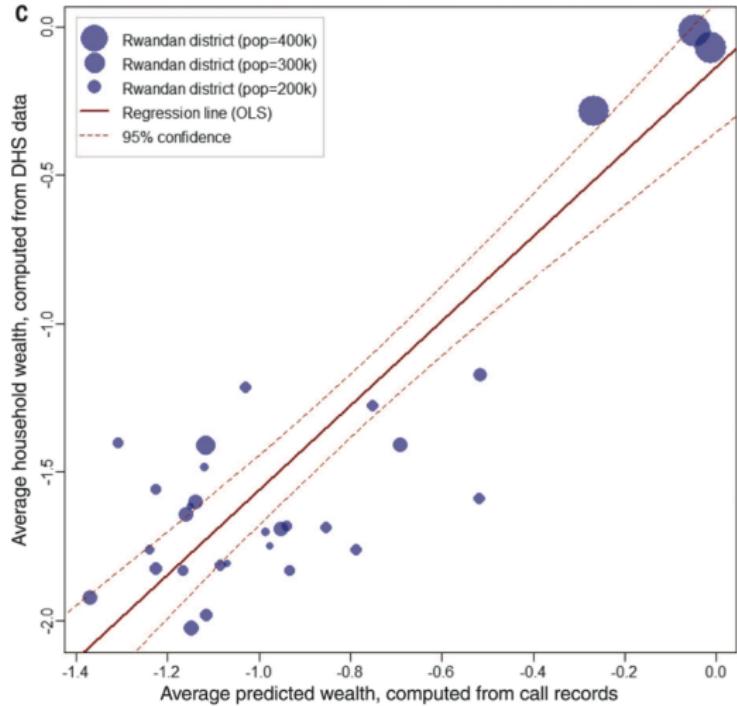






**A**





- ▶ 10 times faster
- ▶ 50 times cheaper

- ▶ *computational and social science*

- ▶ *computational and social science*
- ▶ often involves ethical/privacy questions that are now considered complex

- ▶ *computational and social science*
- ▶ often involves ethical/privacy questions that are now considered complex
- ▶ combines readymades and customades







Readymades



Custommades

[https://commons.wikimedia.org/wiki/File:Duchamp\\_Fountainne.jpg](https://commons.wikimedia.org/wiki/File:Duchamp_Fountainne.jpg)

[https://commons.wikimedia.org/wiki/File:%27David%27\\_by\\_Michelangelo\\_JBU0001.JPG](https://commons.wikimedia.org/wiki/File:%27David%27_by_Michelangelo_JBU0001.JPG)

- ▶ *computational* and *social science*
- ▶ often involves ethical/privacy questions that are now considered complex
- ▶ combines readymades and custommades

- ▶ *computational and social science*
- ▶ often involves ethical/privacy questions that are now considered complex
- ▶ combines readymades and custommades
- ▶ involves five key communities: social science, data science, business people, privacy advocates, policy makers

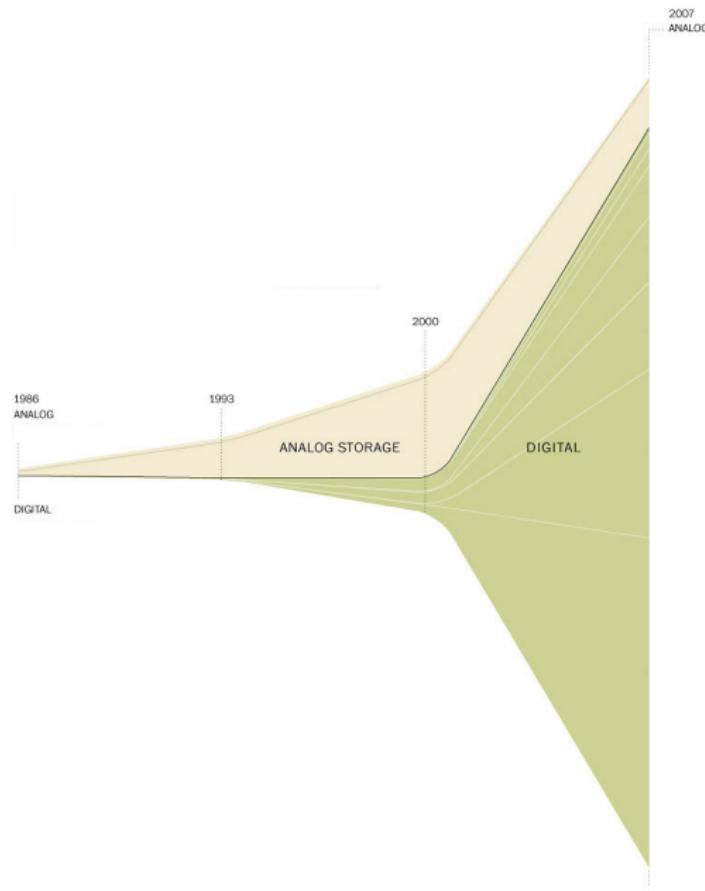
Balance between social science, data science, business people,  
privacy advocates, and policy makers

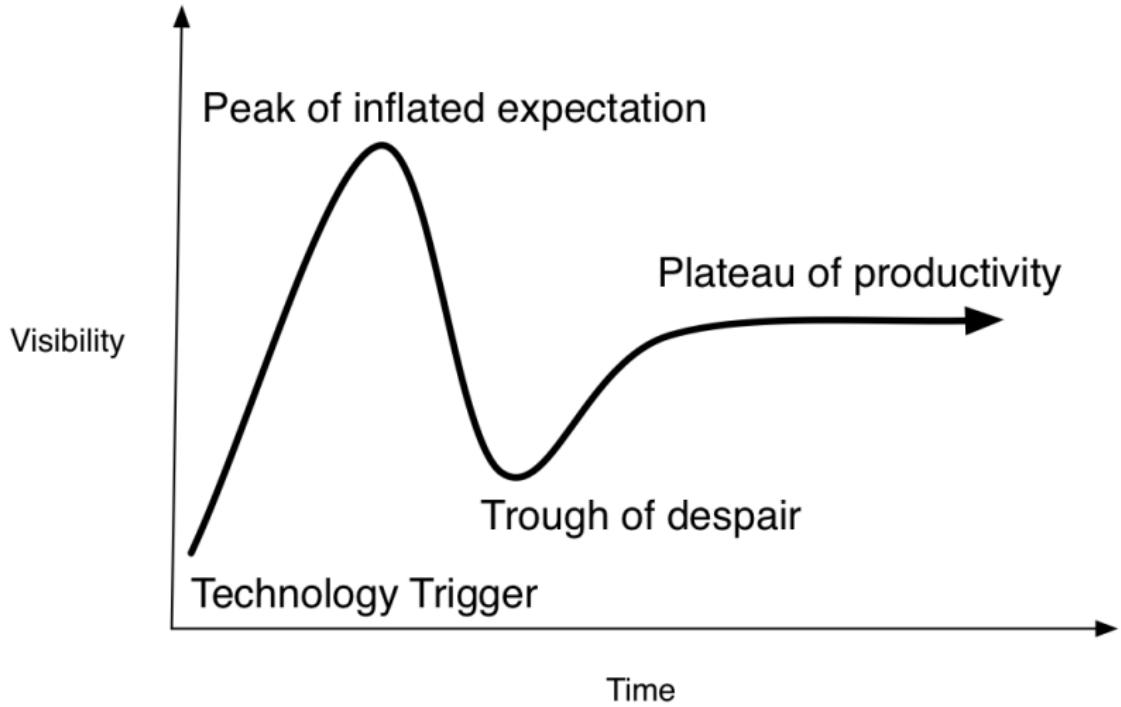
Health Ecosystem vs Horrific Monoculture

# What is computational social science?

- ▶ *computational and social science*
- ▶ often involves ethical/privacy questions that are now considered complex
- ▶ combines readymades and custommades
- ▶ involves five key communities: social science, data science, business people, privacy advocates, policy makers

Isn't computational social science a fad?





[https://commons.wikimedia.org/wiki/File:Gartner\\_Hype\\_Cycle.svg](https://commons.wikimedia.org/wiki/File:Gartner_Hype_Cycle.svg)

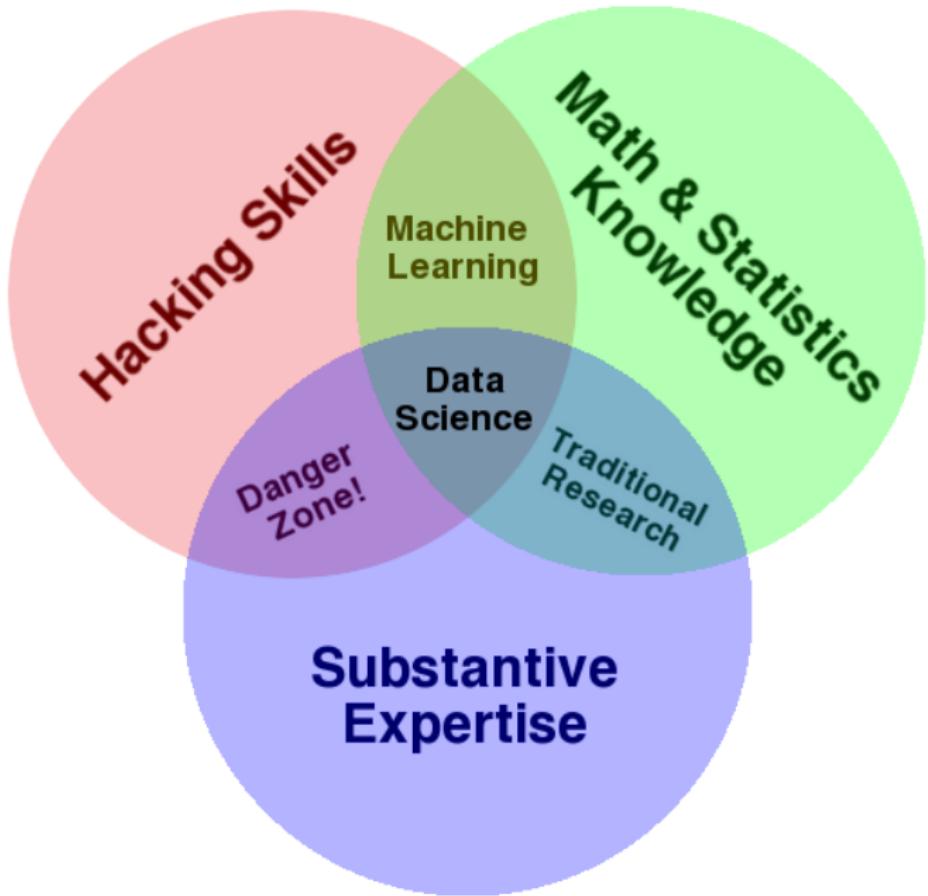
Isn't computational social science a fad?

No

How do we create computational social science, individually and as a community?

Social Scientists ←→ Data Scientists

# What is data science?



# What is data science?

Academic data science and industrial data science

Academic data science (think of Donoho article)

- ▶ goes back to Tukey (1962)

Academic data science (think of Donoho article)

- ▶ goes back to Tukey (1962)
- ▶ learning from data.

## Academic data science (think of Donoho article)

- ▶ goes back to Tukey (1962)
- ▶ learning from data.
- ▶ it includes statistics and much more

Industrial data science (think of Provost and Fawcett)

- ▶ lots of engineering

## Industrial data science (think of Provost and Fawcett)

- ▶ lots of engineering
- ▶ certain kinds of problems—churn and pop-tarts—which are not as far from social science as they appear



Why take ideas from and contribute to  
academic and industrial data science?

Data science alone is not enough if we want to study social behavior (e.g., Anderson)

Wallach: I don't get it. Why is that research?

Wallach: I don't get it. Why is that research?

Computer science	Social science
Study anything	Study social things
Methods driven	Question driven
Large found data	Small designed data
Prediction	Explanation

Another important difference between social science:



Social Scientists ←→ Data Scientists

*Computational social science* vs Computational *social science*

My goals for this module:

- ▶ SWBAT *describe* computational social science
- ▶ SWBAT assess whether computational social science is a fad
- ▶ SWBAT *describe* the intellectual traditions that underly computational social science