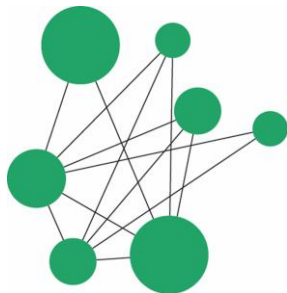
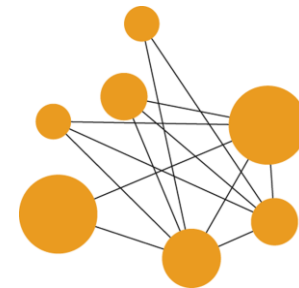


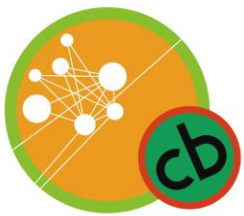


# Ethics of Data Science

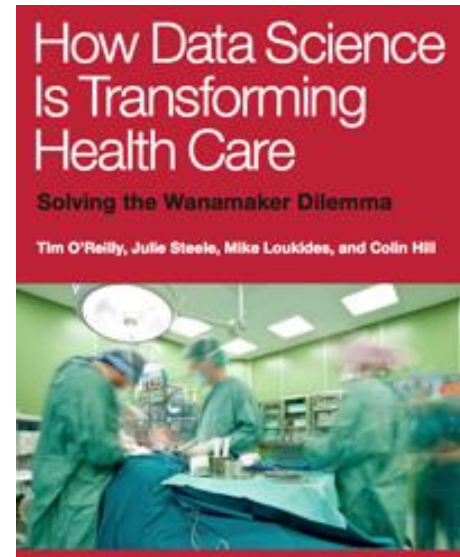
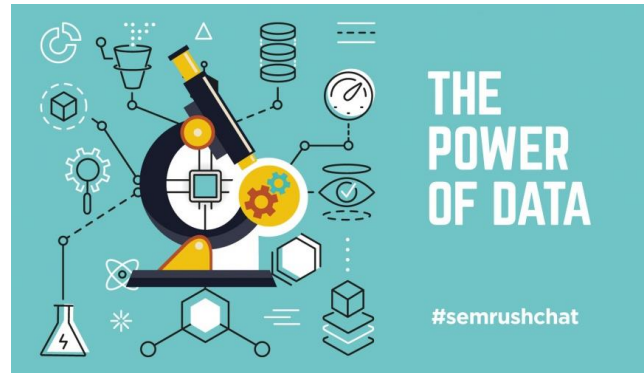
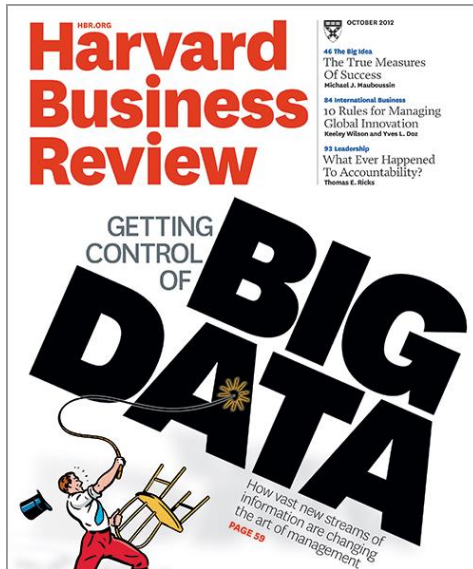
Lawrence Hunter, Ph.D.  
Director, Computational Bioscience Program  
University of Colorado School of Medicine

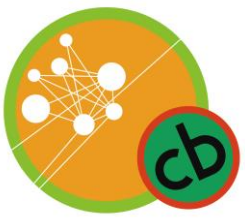


[Larry.Hunter@ucdenver.edu](mailto:Larry.Hunter@ucdenver.edu)  
<http://compbio.ucdenver.edu/Hunter>



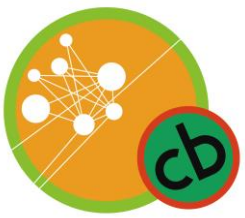
# “Data Science” is everywhere





# What is Data Science?

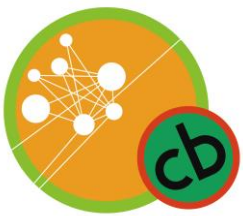
- *Machine Learning*: Data-driven model selection (through a large space of possible models)
- “I think data-scientist is a sexed up term for a statistician,” — *Nate Silver*
- “Our data science team brings together three things: statistics, programming, and product knowledge.” — *Brad Schumitsch, Amazon/Twitch*



# Why this could be good

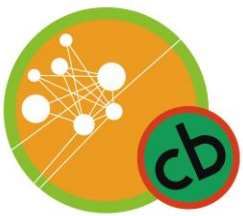
- Algorithms for tasks that were not previously amenable to automation (e.g. image analysis)
- Advantages over humans doing similar tasks:
  - Inexpensive/scalable/fast
  - Consistent and verifiable
  - More accurate\*
  - More fair\*, less subject to social biases

\* Maybe. Sometimes.



# Why this could be bad

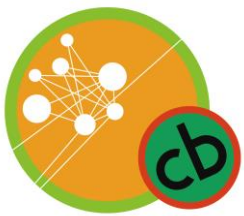
- Data are (about) people, can cause harm
- Algorithmic outcomes often not explainable
- A lot of data incidentally produced by daily life:
  - Social media
  - Ubiquitous cameras, microphones, location tracking
  - Medical treatment
- Important new uses
  - **Legal**: Surveillance, “predictive” policing, sentencing, fraud detection, military applications
  - **Economic**: School admissions, hiring/promotion, loans, insurance, accounting controls, advertising
  - **Medical**: Health insurance, diagnosis, decision making



# Some ethical concerns

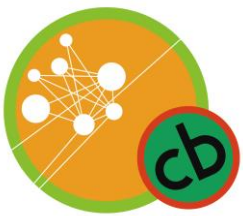
- Preserving privacy
  - Methods for handling sensitive data
  - Uses of data science that undermine privacy
- Avoiding bias
  - Data selection and unintentional red-lining
  - Re-inscription of existing biases
- Mitigating malicious attacks
  - Intentional subversion of machine learning systems
  - Hazards of learning from the open internet





# Anonymized data isn't always

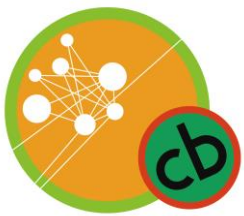
- In 1997, Latanya Sweeney identified the Governor of Massachusetts' medical records.
  - Massachusetts released hospital records anonymized by removing names, addresses and SSNs
  - Voter records have name, address, ZIP code, birth date, and sex of every voter
  - Sweeney used zip code, birthdate and gender to uniquely identify Weld's records
  - 87% of US identified by zip, birthdate & gender
- Similar with Netflix (using IMDB) and search logs



# Privacy $\neq$ Security

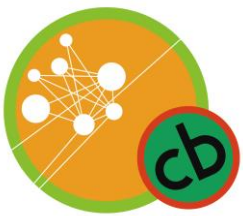
- **Some data cannot be anonymized**
  - Genome sequences are inherently identifying
  - Even a few hundred well-picked SNPs...
- **Often, people's desires about their data involve questions of *trust***
  - Willingness to share medical data with academic researchers, but not pharmaceutical companies
- **Privacy is not a binary value**
  - Different sorts of exposure to different sorts of people evoke different responses





# Privacy and technology

- Privacy preserving technologies:
  - K-anonymity
  - Ignorant processing
- Privacy invading technologies:
  - Identifying people and their locations by cell-phone metadata
  - Descrambling pixelated images: “Defeating Image Obfuscation with Deep Learning”  
McPhearson, et al. 2016



# Data Sharing

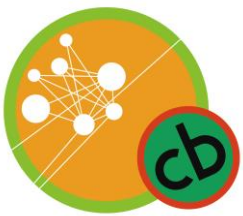
- Data sharing can be of great scientific value
- Often, data generators control (no sharing)
- New models emerging requiring more sharing
  - Genomics / sequences
  - Large NIH grants
  - Clinical trials?
- Participants are surprised it doesn't happen

*NATURE GENETICS* | CORRESPONDENCE

Celebrating parasites

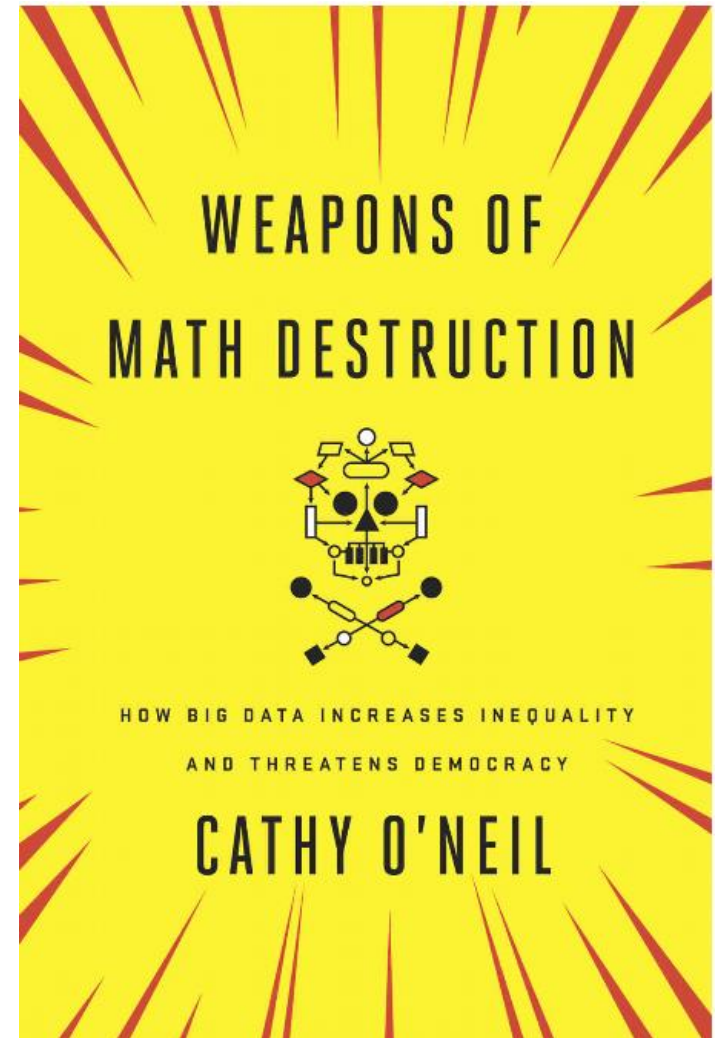
**The SPRINT Data Analysis Challenge** ⓘ

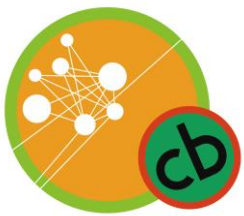
To explore the potential of clinical trial data sharing, the New England Journal of Medicine (NEJM) is hosting a challenge: use the data underlying a recent NEJM article to identify a novel clinical finding that advances medical science.



# Data Science and Bias

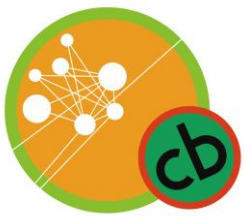
- “Objective” algorithms are thought to be free of the biases that plague people.
- Algorithms, especially ones that learn, can inadvertently re-inscribe those biases
- Algorithms are opaque, hard to interrogate
- Increasingly widespread





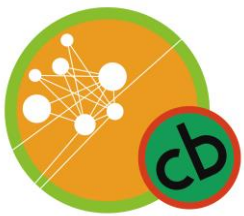
# Discrimination and its proxies

- Illegal, and generally perceived as wrong to make choices based on race, gender, religion, national origin, etc.
- However, proxies for these are everywhere:
  - Zip codes
  - Names (gender, race, national origin)
  - Purchase histories (including movies or tv shows)
- Machine learning that uses biased historical record + any proxy is likely to re-inscribe bias



# Discrimination in Online Ad Delivery

- Sweeney observed in 2013 that black-identifying names turned out to be much more likely than white-identifying names to generate ads that including the word “arrest” (60 per cent versus 48 per cent).
- Google uses a learning algorithm to place ads that are most often clicked on.
- Likely to be a reflection of people clicking on those ads more for ‘black’ names



# Adversarial environments

- Since there is a lot riding on algorithms, people have an interest in manipulating them.
- Many effective strategies

## Practical Black-Box Attacks against Machine Learning

Nicolas Papernot, Patrick McDaniel, Ian Goodfellow, Somesh Jha, Z. Berkay Celik, Ananthram Swami

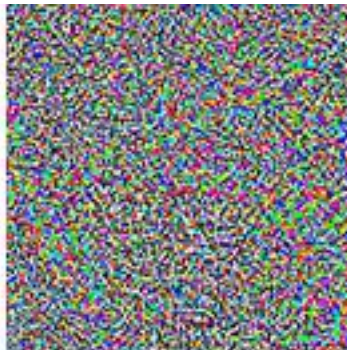
(Submitted on 8 Feb 2016 (v1), last revised 19 Mar 2017 (this version, v4))



"panda"

57.7% confidence

+  $\epsilon$

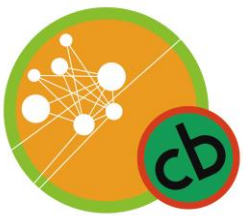


=



"gibbon"

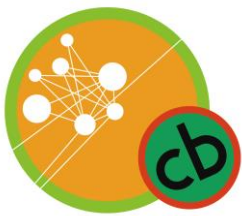
99.3% confidence




# Beware the open internet

- “Tay” was a chatbot designed last year by Microsoft to interact with people over Twitter
  - Built by "mining relevant public data" and combining that with input from editorial staff, "including improvisational comedians." The bot is supposed to learn and improve as it interacts with users
- Within 24 hours of being unveiled, it was pulled after making many racist, sexist, etc. statements
  - “As it learns, some of its responses are inappropriate and indicative of the types of interactions some people are having with it. We're making some adjustments to Tay.”



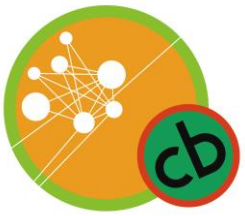


# Ten simple rules for responsible big data research

Matthew Zook , Solon Barocas, danah boyd, Kate Crawford, Emily Keller, Seeta Peña Gangadharan, Alyssa Goodman, Rachelle Hollander, Barbara A. Koenig, Jacob Metcalf, Arvind Narayanan, Alondra Nelson, Frank Pasquale

Published: March 30, 2017 • <https://doi.org/10.1371/journal.pcbi.1005399>

1. Acknowledge that data are people and can do harm
2. Recognize that privacy is more than a binary value
3. Guard against the re-identification of your data
4. Practice ethical data sharing
5. Consider the strengths and limitations of your data; big does not automatically mean better
6. Debate the tough, ethical choices
7. Develop a code of conduct for your organization, research community, or industry
8. Design your data and systems for auditability
9. Engage with the broader consequences of data and analysis practices
10. Know when to break these rules



Let's discuss