

# Privacy for Data Analysis and ML

CS848 Fall 2024



UNIVERSITY OF  
**WATERLOO**



# Instructor



## Xi He:

- Research interest: privacy and security for data management and analysis
- CS848, Fall 2024:
  - Thur: 1:00pm – 3:40pm (DC2568)



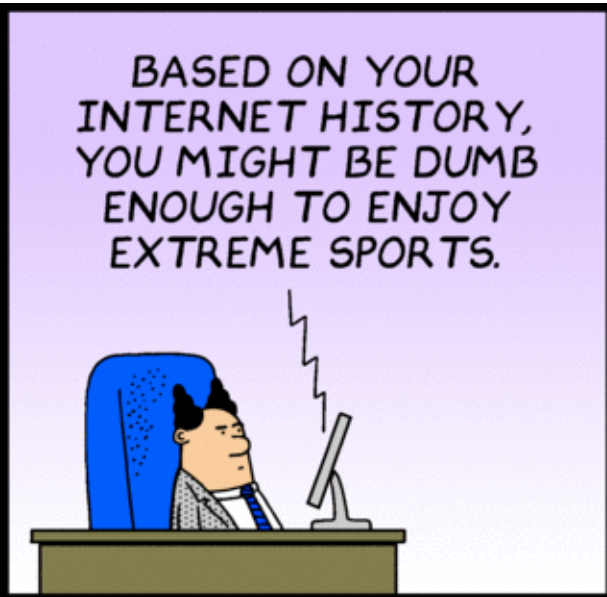
## Bailey Kacsmar:

- University of Alberta
- Research interest: human-centered technical privacy solutions
- Co-designer and guest lecturer

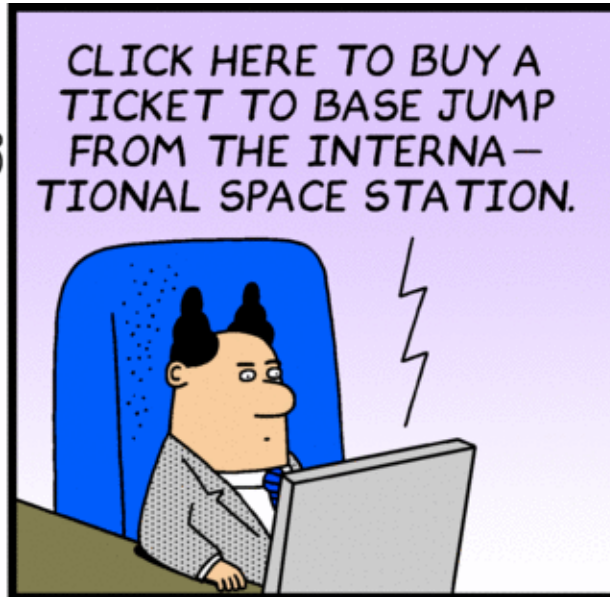
Tell me ...

... why do you want to do this course?

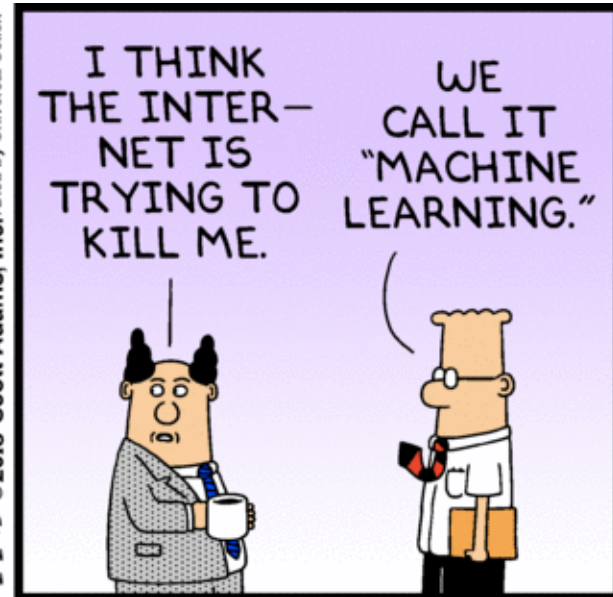
# Personalization ...



Dilbert.com DilbertCartoonist@gmail.com

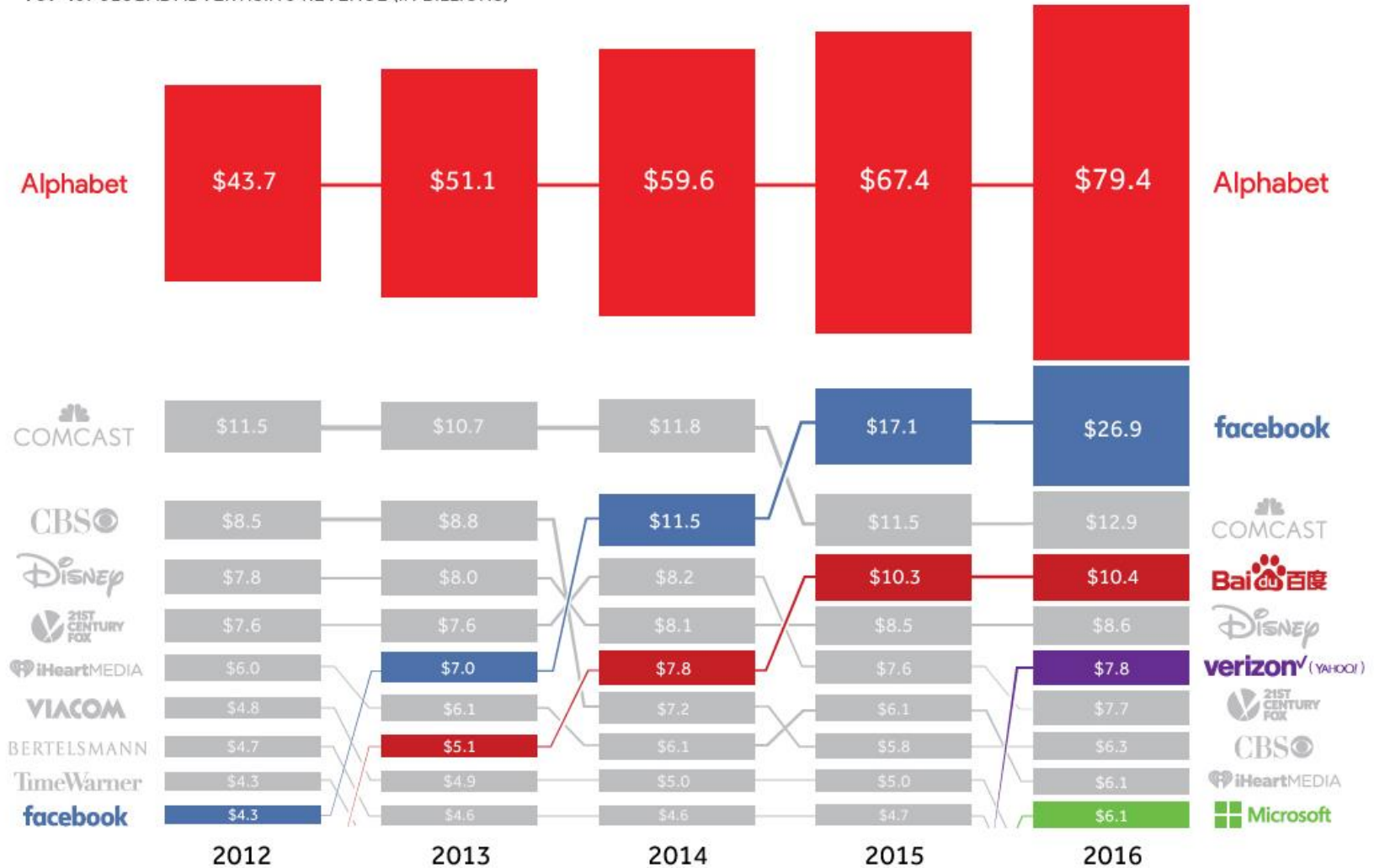


2-2-13 © 2013 Scott Adams, Inc. / Dist. by Universal Uclick



# Online Advertising

TOP 10: GLOBAL ADVERTISING REVENUE (IN BILLIONS)



SOURCE: Bloomberg, Zenith Media



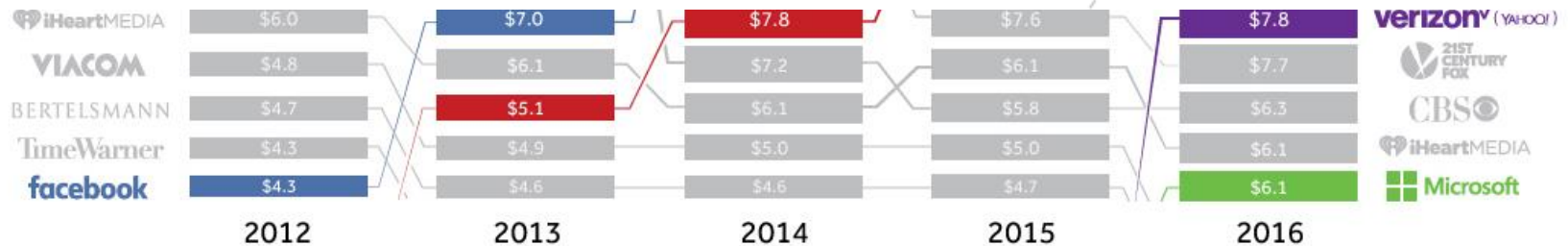
# Online Advertising

TOP 10: GLOBAL ADVERTISING REVENUE (IN BILLIONS)



**Ad-Supported Internet Brings Over \$1 Trillion To The U.S. Economy, Representing 6 Percent Of Country's Total GDP, According To IAB Study Led By Harvard Business School Professor**

03.15.17



# TAPESTRY SEGMENTATION

The Fabric of America's Neighborhoods



## UNITED STATES OF AMERICA

Total Population: 312,688,000 Median Income: \$31,000 Home Diversity Index: 61%  
 Total Households: 128,979,000 Median Net Worth: \$17,000 Average Household Size: 2.38  
 Median Age: 35.4 Diversity Index: 62.2 Home Value: \$177,000



### LIFESTYLE SUMMARY GROUPS

- 1. Affluent, Uppscale, Uptown, and Family Landscapes
- 2. Gen X Urban
- 3. Cozy Country Living
- 4. Ethnic Enclaves
- 5. Middle Country
- 6. Senior Styles
- 7. Rustic Outposts
- 8. Midtown Singles
- 9. Hometown
- 10. Scholars and Patriots

### URBANIZATION SUMMARY GROUPS

- 1. Affluent, Uppscale, Uptown, and Family Landscapes
- 2. Gen X Urban
- 3. Cozy Country Living
- 4. Ethnic Enclaves
- 5. Middle Country
- 6. Senior Styles
- 7. Rustic Outposts
- 8. Midtown Singles
- 9. Hometown
- 10. Scholars and Patriots

### DEFINITIONS BY THE SEGMENT DESCRIPTIONS

Household (HH) equal family and non-family

- 1 Family Income range
- 2 Family Income range
- 3 Family Income range
- 4 Family Income range
- 5 Family Income range
- 6 Family Income range
- 7 Family Income range
- 8 Family Income range
- 9 Family Income range
- 10 Family Income range

FOR MORE INFORMATION ABOUT TAPESTRY SEGMENTATION  
 Visit [esri.com](http://esri.com)  
 or call 1-800-440-4444



**SEGMENT LEGEND**

1 Segment Name  
 2 Household Income Range  
 3 Family Income Range  
 4 Family Income Range  
 5 Family Income Range  
 6 Family Income Range  
 7 Family Income Range  
 8 Family Income Range  
 9 Family Income Range  
 10 Family Income Range

10-point Income Scale: 100% (top) to 0% (bottom)

10-point Education Scale: 100% (top) to 0% (bottom)

10-point Net Worth Scale: 100% (top) to 0% (bottom)

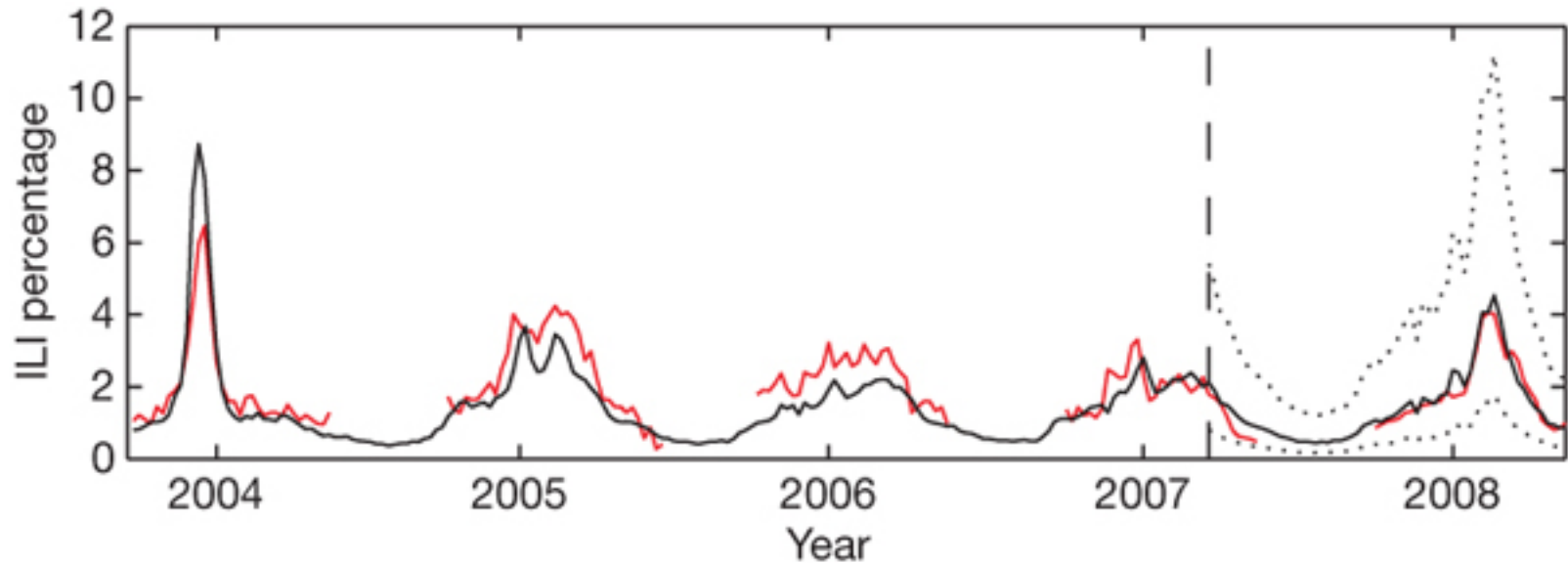
10-point Diversity Index Scale: 100% (top) to 0% (bottom)

10-point Mobility Index Scale: 100% (top) to 0% (bottom)

10-point Density Index Scale: 100% (top) to 0% (bottom)

10-point Growth Index Scale: 100% (top) to 0% (bottom)

# Health



**Red:** official numbers from Center for Disease Control and Prevention; weekly  
**Black:** based on Google search logs; daily (potentially instantaneously)

## Detecting influenza epidemics using search engine query data

<http://www.nature.com/nature/journal/v457/n7232/full/nature07634.html>



# IMPRECISION MEDICINE

For every person they do help (blue), the ten highest-grossing drugs in the United States fail to improve the conditions of between 3 and 24 people (red).

## 1. ABILIFY (aripiprazole)

Schizophrenia



## 2. NEXIUM (esomeprazole)

Heartburn



## 3. HUMIRA (adalimumab)

Arthritis

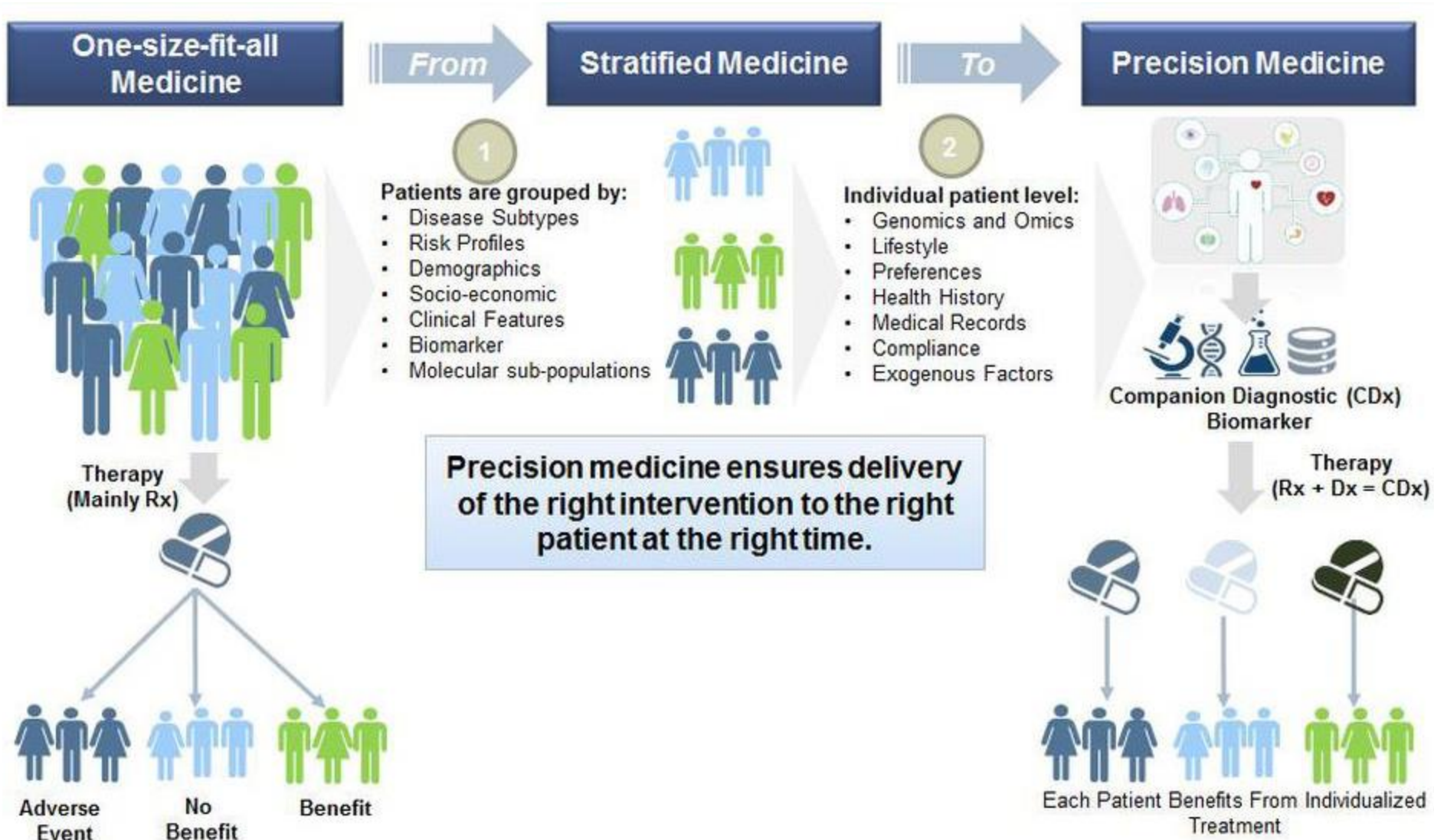


## 4. CRESTOR (rosuvastatin)

High cholesterol



# Precision Medicine



# Predictive Policing

How can predictive policing drive proactive crime prevention?

IBM



## Manchester Police Department

Protects and serves the 110,000 citizens of Manchester, New Hampshire



Needed a smarter way to decide where its 237 officers should patrol



Worked with Ironside to harness IBM® SPSS® Modeler software to help predict where crimes were likely to occur

12%

reduction in robberies

21%

reduction in burglaries

32%

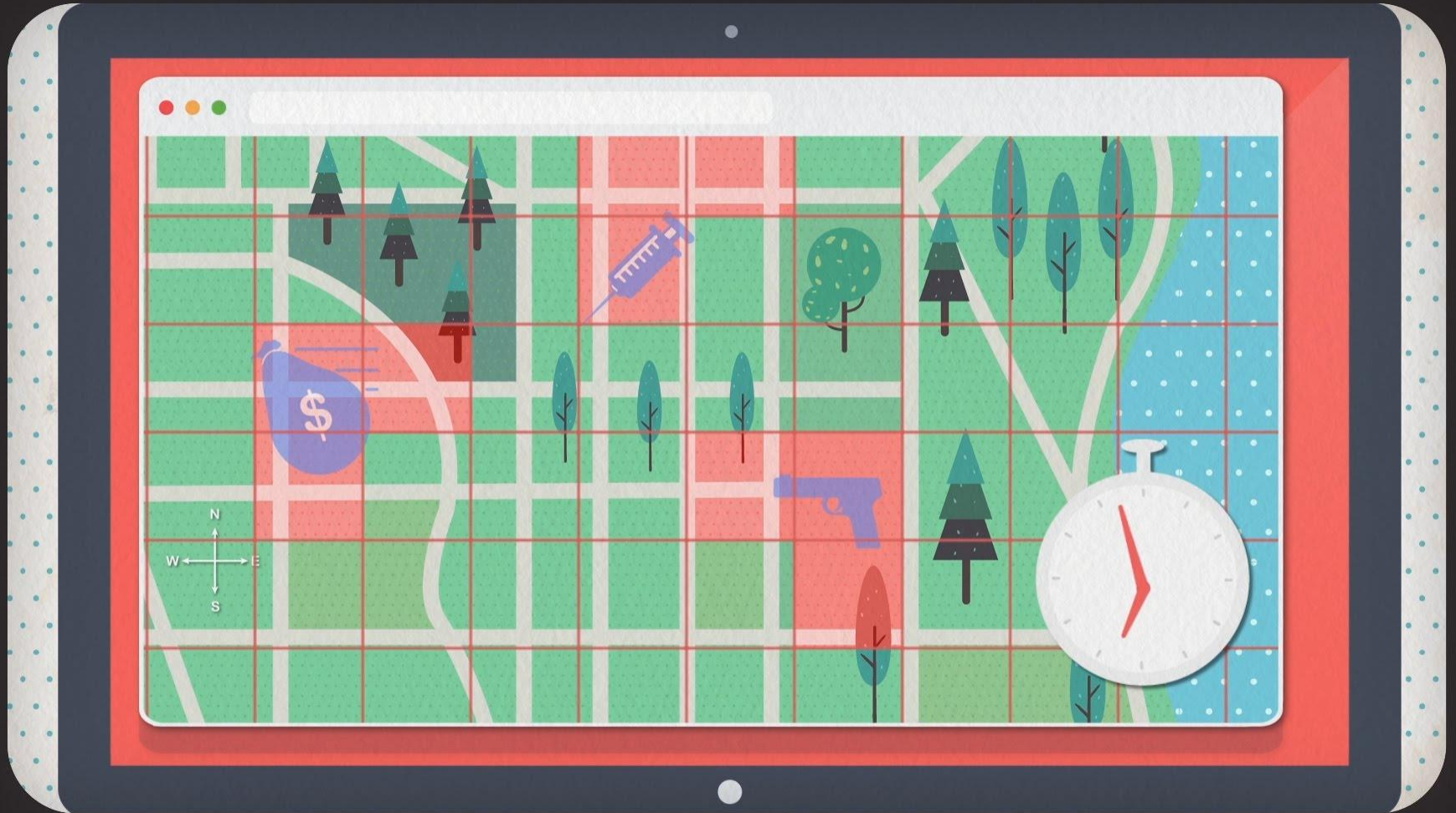
reduction in thefts from vehicles



Entire Manchester police force was monitored and these numbers represent a comparison of statistics from July 8, 2015 through December 8, 2015 compared to the same six month period (July - December) in 2014.

© IBM Corporation 2015. DocNumber:TRC-US26-00. Sources: <http://ibm.co.uk/MAA26> | <http://ibm.co.uk/Lawen1>

# Predictive Policing



# The dark side of the force...



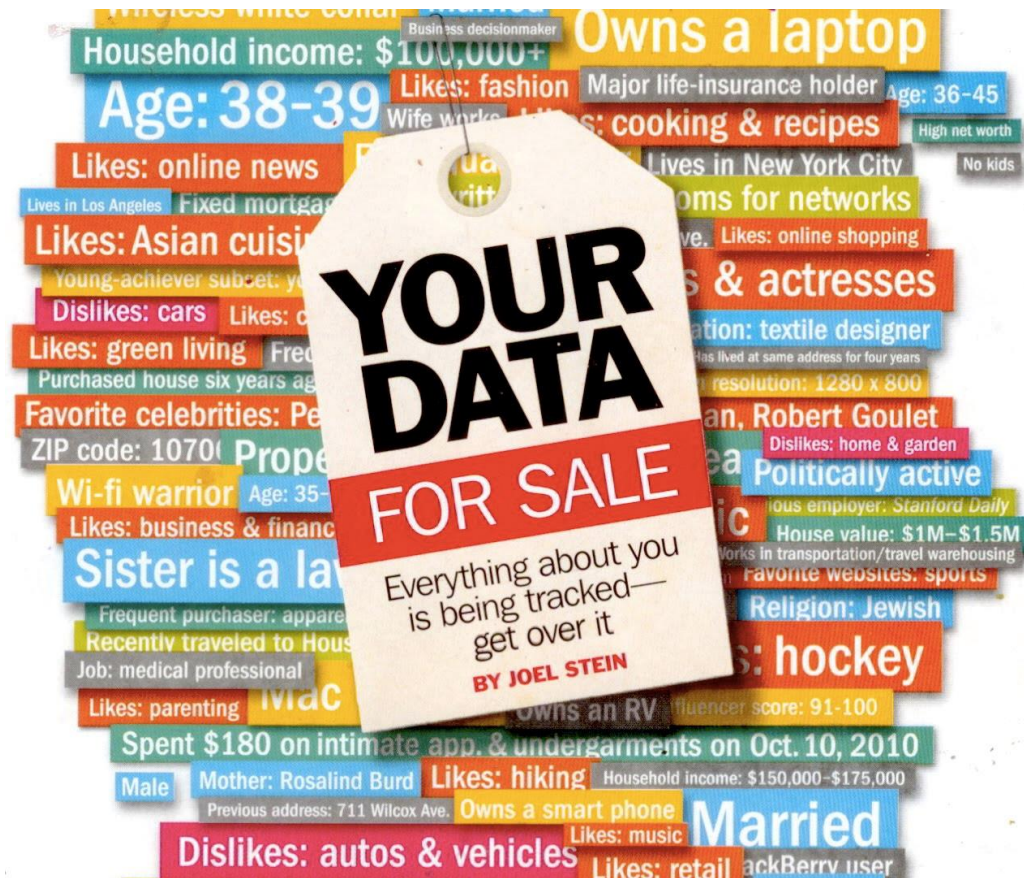
# 39% of the experts agree...

Thanks to many changes, including the building of “the Internet of Things,” human and machine analysis of **Big Data will cause more problems than it solves** by 2020. The existence of huge data sets for analysis will **engender false confidence in our predictive powers** and will lead many to make **significant and hurtful mistakes**. Moreover, analysis of Big Data will be **misused by powerful people and institutions with selfish agendas** who manipulate findings to make the case for what they want. And the advent of Big Data has a harmful impact because **it serves the majority (at times inaccurately) while diminishing the minority** and ignoring important outliers. Overall, the rise of Big Data is a big negative for society in nearly all respects.

— 2012 Pew Research Center Report

<http://pewinternet.org/Reports/2012/Future-of-Big-Data/Overview.aspx>

# Where is the data coming from?



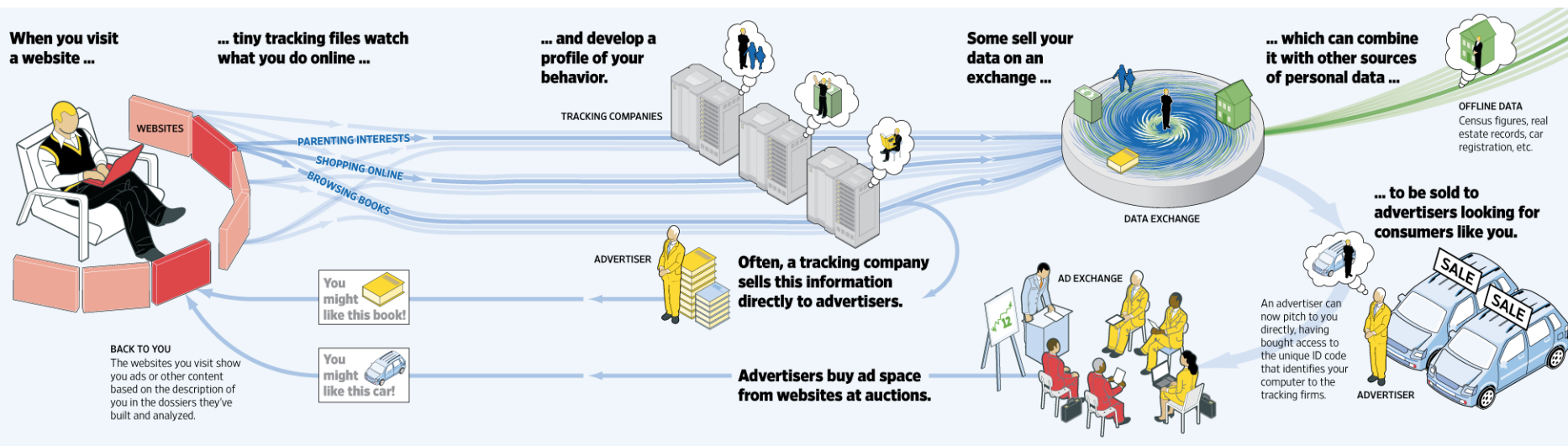
# Where is the data coming from?

- Census surveys
- IRS Records
- Photos
- Videos
- Medical records
- Insurance records
- Smart phone Sensors
- Mobility trajectories
- Search logs
- ...
- Browse logs
- Shopping histories

**Very sensitive information ...**



# How is this data collected?



<http://graphicsweb.wsj.com/documents/divSlider/media/ecosystem100730.png>

# Isn't my data anonymous ?



# Device Fingerprinting

A typical computer broadcasts hundreds of details about itself when a Web browser connects to the Internet. Companies tracking people online can use those details to 'fingerprint' browsers and follow their users.

**Timestamp** One fingerprinting technique compares the time on a person's computer to the time on a Web server down to the millisecond.

**User ID** Once a device has been fingerprinted, it is assigned a 'token,' or ID number, that can be used to track a user's online activities.

one: 300

... / (h:mm:ss.ms)  
 / (+1:59:59.560)  
 / (+1:59:59.548)

...: Stainless  
 ...ic, Stain

...isplay Light,  
 ...ronic Display B  
 ...e Disp Cond Semi,  
 ...e Disp Cond, Chron  
 ...d Light, Chronicl  
 ...ronic Disp  
 ...p Comp, C

... Settings,  
 ... (via Flash)

Screen size and  
 1280x1024x32

Browser Plug-in De  
 Adobe Acrobat; A  
 ...32.dll

Device Token: 28AB-ECDD-7A8C-3D7A-2563-AE87-C551-5D4D

...ent in Web  
 ...formation, vis  
 ...REF=http://www.a  
 ...e/>QuickTime</A>  
 ...npqtplugin.dll; (S  
 ...SDP stream descr  
 ...application/x-sd  
 ...am descri

...ation/itunc  
 ...r Agent: Mozilla/  
 ...Windows NT 5.1; er  
 ...pleWebKit/534.10 (C  
 ...cko) Chrome/8.0.55  
 ...ari/534.10

**Fonts** Not all machines have the same typefaces installed. The order the fonts were installed can also distinguish one computer from another.

**Screen Size** Things like the size of the screen and its color settings can help websites display content correctly, but also can be used to identify machines.

**Browser Plugins** The mix of QuickTime, Flash and other 'plugins' (small pieces of optional software within a browser) can vary widely.

**User Agent** This is tech-speak for the type of Web-browsing software used. It can include specific details about the computer's operating system, too.

# PANOPTICCLICK<sup>3.0</sup>

Is your browser safe against tracking?

Your browser fingerprint **appears to be unique** among the 2,050,572 tested in the past 45 days.

Currently, we estimate that your browser has a fingerprint that conveys **at least 20.97 bits** of identifying information.

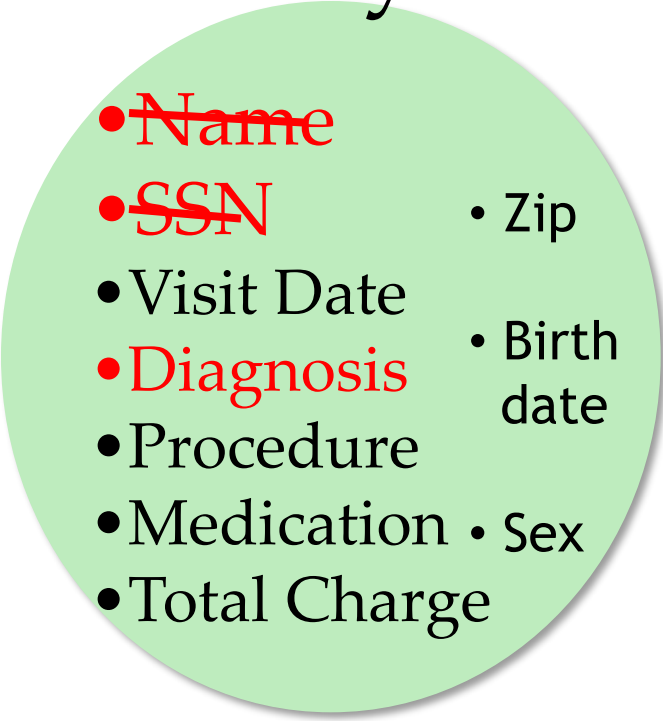
<https://panopticclick.eff.org/>

Let's get rid of unique identifiers ...



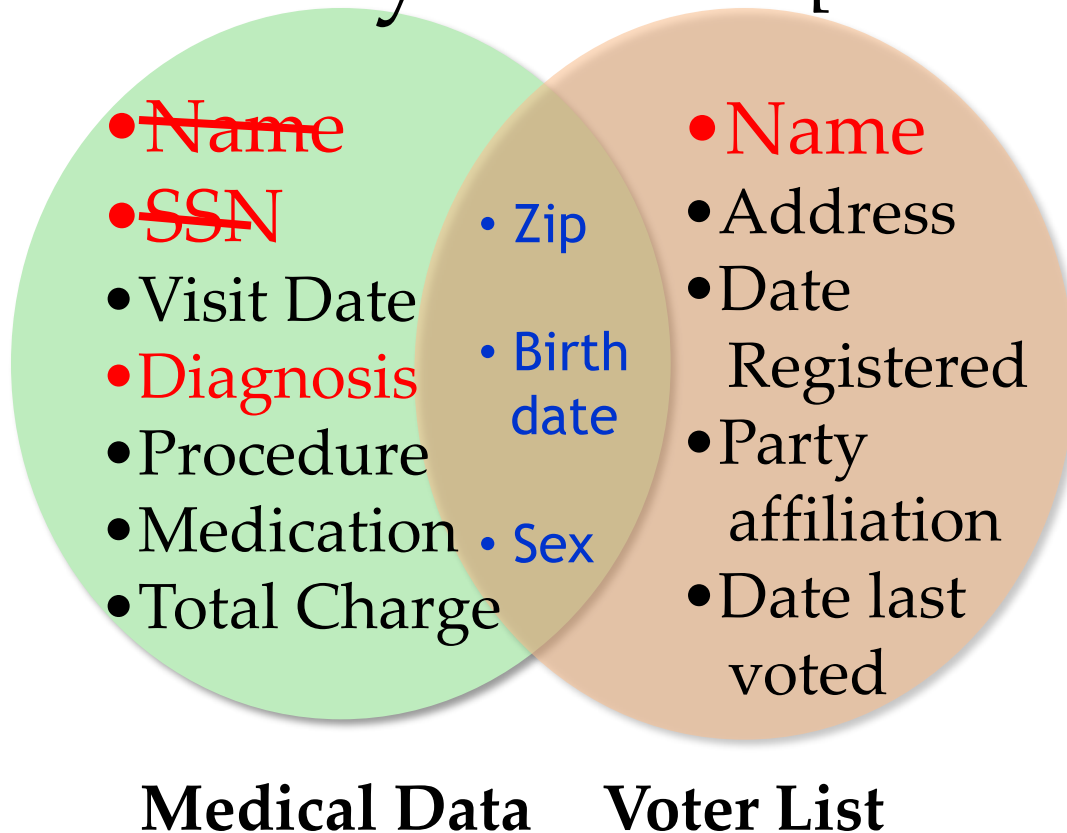
**HIPAA**  
COMPLIANT

# The Massachusetts Governor Privacy Breach [Sweeney IJUFKS 2002]

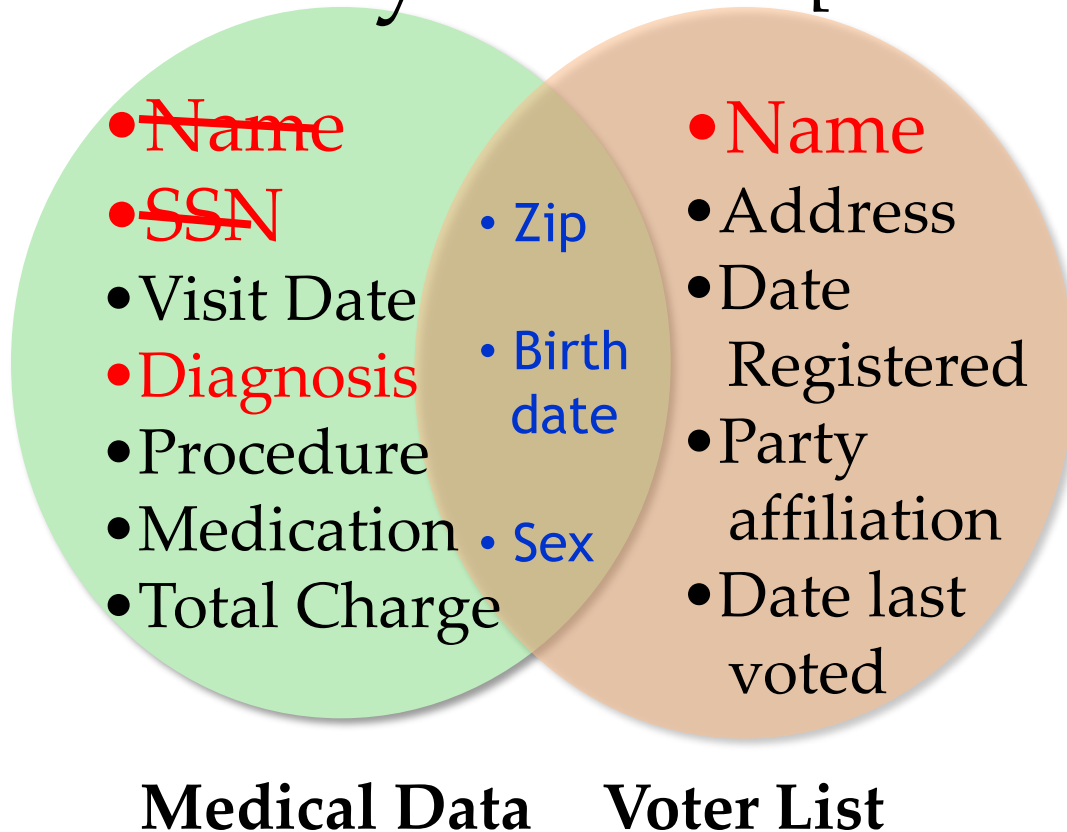
- 
- ~~Name~~
  - ~~SSN~~
  - Visit Date
  - ~~Diagnosis~~
  - Procedure
  - Medication
  - Total Charge
  - Zip
  - Birth date
  - Sex

**Medical Data**

# The Massachusetts Governor Privacy Breach [Sweeney IJUFKS 2002]



# The Massachusetts Governor Privacy Breach [Sweeney IJUFKS 2002]

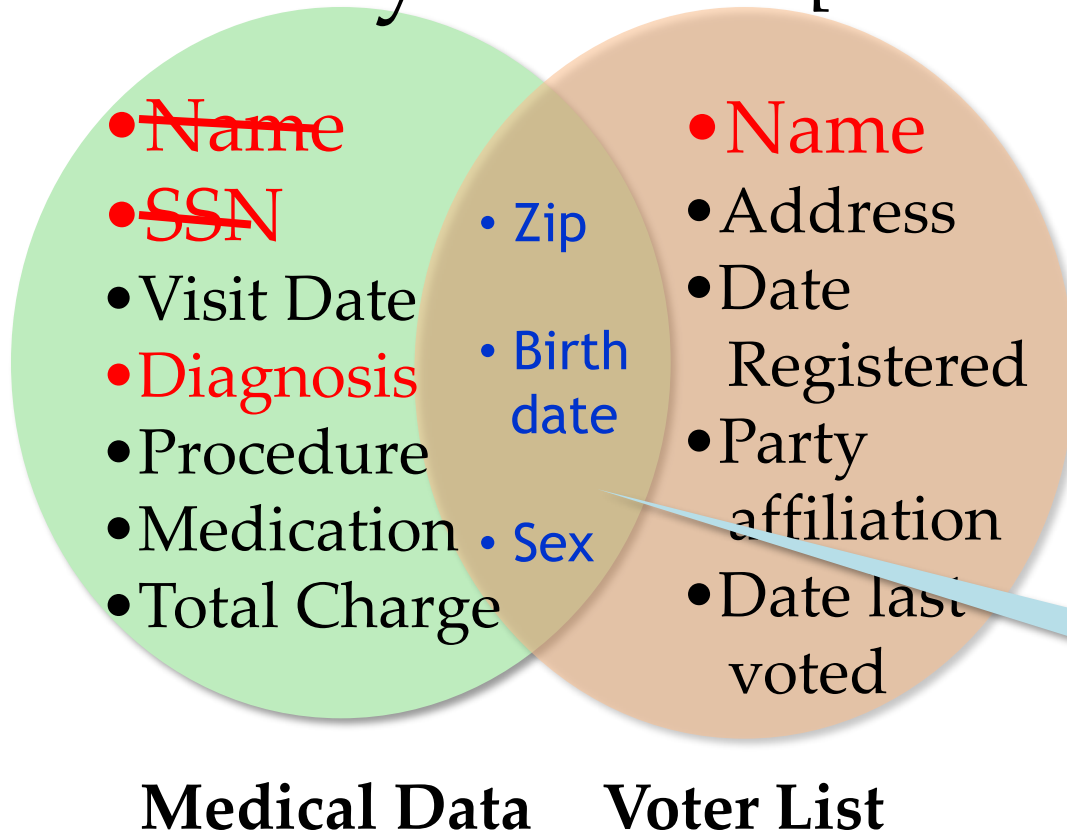


- **Governor of MA uniquely identified using ZipCode, Birth Date, and Sex.**

**Name linked to  
Diagnosis**



# The Massachusetts Governor Privacy Breach [Sweeney IJUFKS 2002]




- 87 % of US population **uniquely identified** using ZipCode, Birth Date, and Sex.

**Quasi Identifier**

# AOL data publishing fiasco

— IN SOLIDARITY WITH THE MANY AOL USERS WHOSE OFTEN EMBARRASSING WEB SEARCHES WERE RELEASED TO THE PUBLIC, I OFFER A SAMPLE OF MY OWN SEARCH HISTORY:



The screenshot shows the Google search interface with the following elements:

- Navigation links:** Web, [Images](#), [Video](#) <sup>New!</sup>, [News](#), [Maps](#), [more »](#)
- Utility links:** [Advanced Search](#), [Preferences](#), [Language Tools](#)
- Search history list:**
  - velociraptors
  - site:imdb.com "jurassic park"
  - raptors
  - dromaeosaurids
  - utahraptor
  - "home depot" deadbolts
  - security home improvement
  - surviving a raptor attack
  - robert bakker paleontologist
  - robert bakker "possible raptor sympathizer"
  - site:en.wikipedia.org surviving a raptor attack
  - learning from mistakes in jurassic park
  - big-game rifles
  - tire irons
  - treating raptor wounds
  - do raptors fear fire
  - how to make a molotov cocktail
  - do raptors fear death
  - can raptors pick locks
  - how to tell if my neighbors are raptors

# AOL data publishing fiasco ...

Xi222	Uefa cup
Xi222	Uefa champions league
Xi222	Champions league final
Xi222	Champions league final 2013
Abel156	exchangeability
Abel156	Proof of deFinetti's theorem
Jane12345	Zombie games
Jane12345	Warcraft
Jane12345	Beatles anthology
Jane12345	Ubuntu breeze
Bob222	Python in thought
Bob222	Entthought Canopy

# User IDs replaced with random numbers

865712345

Uefa cup

865712345

Uefa champions league

865712345

Champions league final

865712345

Champions league final 2013

236712909

exchangeability

236712909

Proof of deFinetti's theorem

112765410

Zombie games

112765410

Warcraft

112765410

Beatles anthology

112765410

Ubuntu breeze

865712345

Python in thought

865712345

Entthought Canopy

# Privacy Breach

[NYTimes 2006]

## A Face Is Exposed for AOL Searcher No. 4417749

By MICHAEL BARBARO and TOM ZELLER Jr.

Published: August 9, 2006

 SIGN IN TO E-  
THIS



# Machine learning models can reveal sensitive information

## Facebook Profile

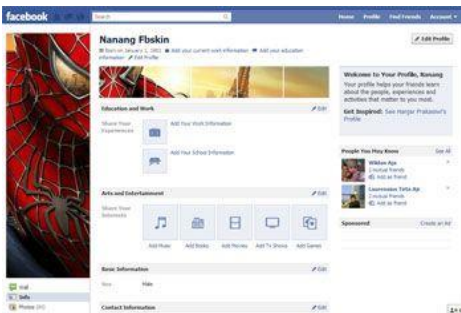
## Number of Impressions

25

+ Who are interested in **Men**

0

+ Who are interested in **Women**



- who live in the **United States**
- who live within 50 miles of **Staten Island, NY**
- between the ages of **23 and 27 inclusive**
- who are **female**
- who are connected to **DogAnd PonyShow**
- in one of the categories: **Pop Culture, Science Fiction/Fantasy, Alternative, Rock, Classic Rock or iPhone**



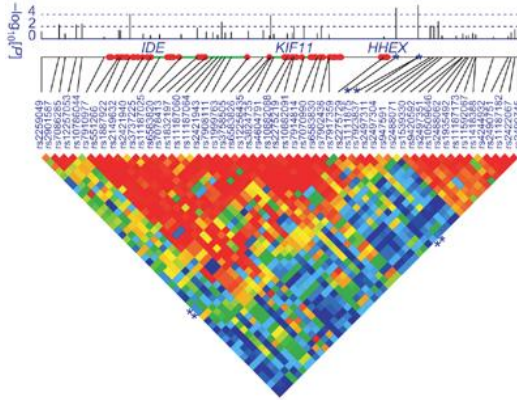
+  
Online Data

Facebook's learning algorithm uses private information to predict match to ad

# Genome wide association studies

[Homer et al PLOS Genetics 08]

Results of a GWAS study



High density SNP profile of Bob



Did Bob participate in the study

**BRACE YOURSELF**

**DEEP LEARNING IS COMING**

[memegenerator.net](http://memegenerator.net)





# Deep Learning

Incredibly powerful tool for ...

- Extracting regularities from data according to a given data
- Amplifying privacy concerns!

Given access to a black-box classifier, can we infer whether a specific example was part of the training dataset?

We can with **shadow training**:

Shokri, R., Stronati, M., Song, C. and Shmatikov, V., 2017, May. **Membership inference attacks against machine learning models**. In *2017 IEEE Symposium on Security and Privacy (SP)*, (pp. 3-18). IEEE.

<i>Dataset</i>	<i>Training Accuracy</i>	<i>Testing Accuracy</i>	<i>Attack Precision</i>
Adult	0.848	0.842	0.503
MNIST	0.984	0.928	0.517
Location	1.000	0.673	0.678
Purchase (2)	0.999	0.984	0.505
Purchase (10)	0.999	0.866	0.550
Purchase (20)	1.000	0.781	0.590
Purchase (50)	1.000	0.693	0.860
Purchase (100)	0.999	0.659	0.935
TX hospital stays	0.668	0.517	0.657

TABLE II: Accuracy of the Google-trained models and the corresponding attack precision.

# This course:

## Learn to combat the dark side



# You will ...

- empirically evaluate privacy
- mathematically formulate privacy
- investigate human-centered privacy
- bridge privacy gaps in policies, practices, and technologies

# Course Format

- Module 1: Empirical privacy
- Module 2: Semantic privacy
- Module 3: Useable privacy
- Module 4: Legal privacy



*Lectures*  
*In-class Exercise*

- Seminars:
  - Paper Reading by Topics



*Read papers*  
*Paper discussion*  
*Research Project*

# Administrivia

- **Website**
  - [https://cs.uwaterloo.ca/~xihe/cs848\\_f24](https://cs.uwaterloo.ca/~xihe/cs848_f24)
  - Schedule (with links to slides, readings, projects, etc.)
- **Grading**
  - Project: 50%
  - Paper reviews, presentation and discussion: 50%
- **LEARN** for submission and grades:
  - <https://learn.uwaterloo.ca/d2l/home/1046490>

# Administrivia - Project

- Projects: (50% of grade)
  - Human centered privacy
  - Privacy attacks (“break” existing privacy algorithms)
  - Privacy-preserving theory/algorithms design
  - Implement/adapt exiting work to new domains
  - Privacy policies and regulations w.r.t. PETs
- Goals:
  - Literature review
  - Some original research/implementation

# Administrivia - Project

- Timeline:
  - Sep 26: Choose Project (ideas will be posted...new ideas welcome)
  - Oct 3: Project proposal (1-4 pages describing the project) **5%**
  - Nov 7: Mid-project review (2-3 page report on progress) **10%**
  - Dec 5 [**TBD**]: Final presentations (10-15 minute talk) **10%**
  - Dec 9: Final report (6-8 page conference style paper) **25%**



# Administrivia - Paper

- Paper presentation and discussion: 50%
  - Paper reviews (15 papers across the term): 15%
  - Seminar style presentations (1-2 per term): 20%
  - Participation in paper discussions: 10%
  - Quality of feedback on peers: 5%
- Details can be found [here](#)

$$\forall i \in [n], d \in \mathcal{S}, \left| \ln \frac{\Pr[T_i \in \mathcal{T} | d_i = d]}{\Pr[T_i \in \mathcal{T} | d_i = \text{NULL}]} \right|$$

$$\left| \frac{\Pr[\text{client}(d) = t]}{\Pr[\text{client}(\text{null}) = t]} \right| \leq \ln \left( \frac{e^\epsilon}{1 + e^\epsilon} \cdot \frac{1 + e^\epsilon}{1} \right) = \epsilon$$

$$\alpha = \frac{3k + 2c_\epsilon \sqrt{\ln(6mk/\beta)}}{\sqrt{n}} = o\left(\frac{\sqrt{\log(p/\beta)}}{\epsilon \sqrt{n}}\right)$$

$$\alpha = \frac{3k + c_\epsilon \sqrt{\ln(4mk/\beta)}}{\sqrt{n}} = o\left(\frac{\sqrt{\log(p/\beta)}}{\epsilon \sqrt{n}}\right)$$

$$\left\{ \left( \frac{v[j] \cdot b[j] + 1}{2} \right), \forall j \in [m] \right\}$$

# What we expect you to know ...

- Strong background in
  - Probability
  - Proof techniques
  
- Some knowledge of
  - Programming with Python
  - Machine learning
  - Statistics
  - Algorithms

# Academic Integrity

- See course website  
[https://cs.uwaterloo.ca/~xihe/cs848\\_f24/](https://cs.uwaterloo.ca/~xihe/cs848_f24/)
- Paper critiques are individual work and submission.
- All suspected cases of violation will be aggressively pursued