Privacy for Data Analysis and ML

CS848 Fall 2024

UNIVERSITY OF WATERLOO | DSg Data Systems Group

Instructor

Xi He:

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

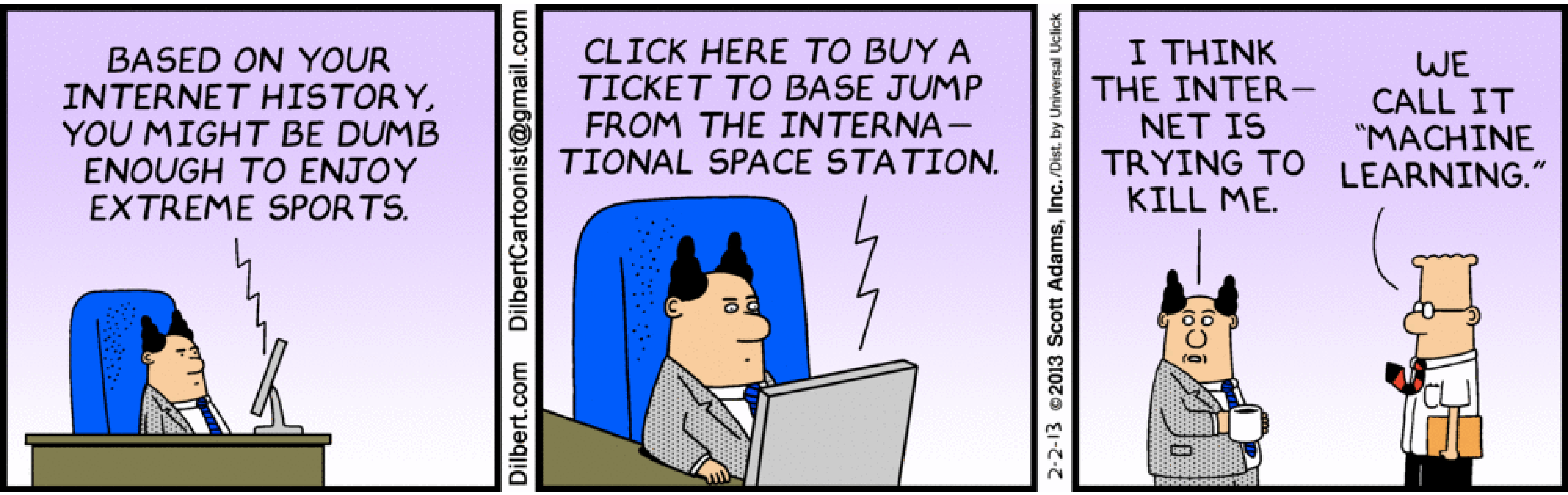
Bailey Kacsmar:

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

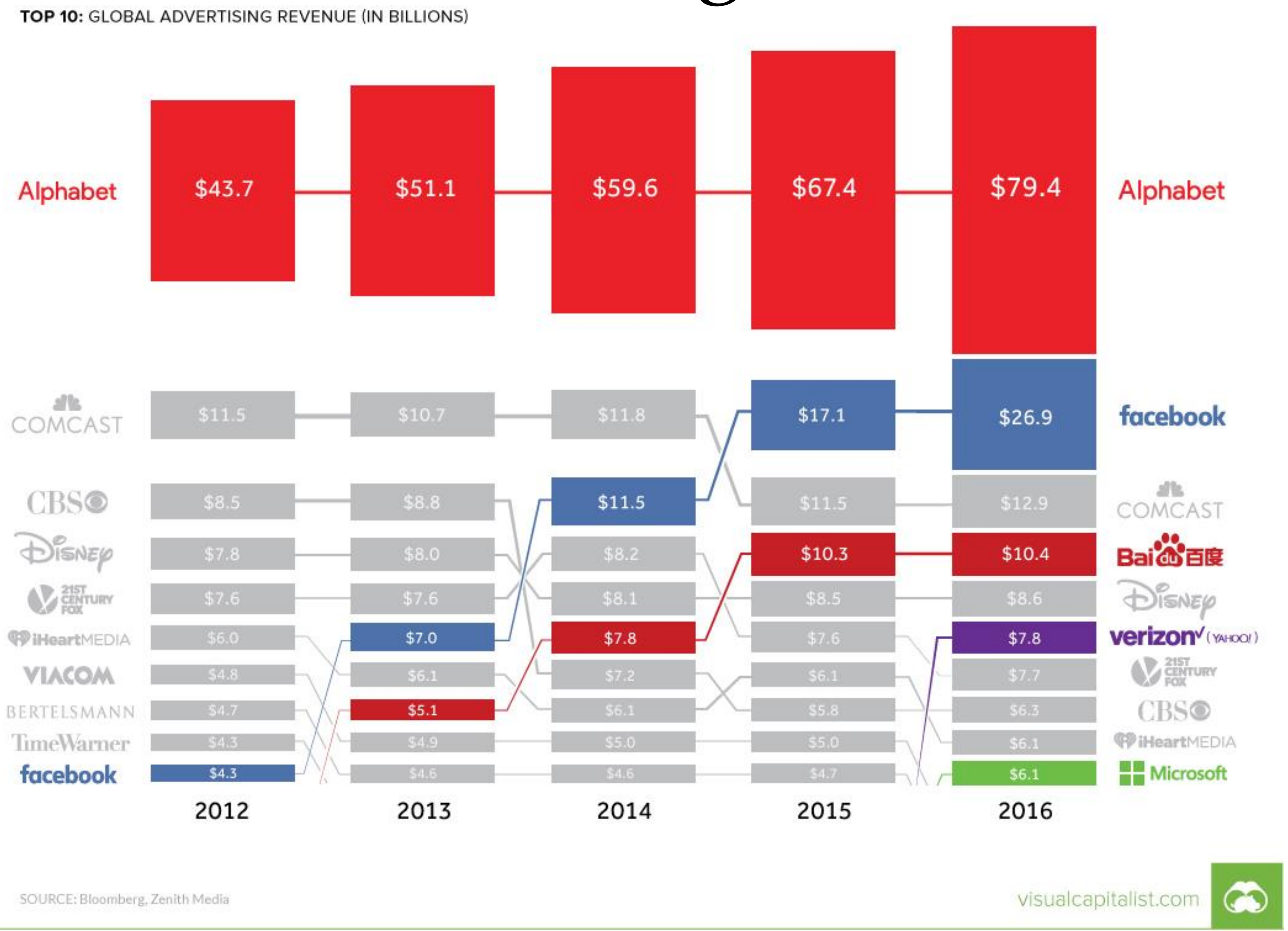
Tell me …

… why do you want to do this course?

Personalization …



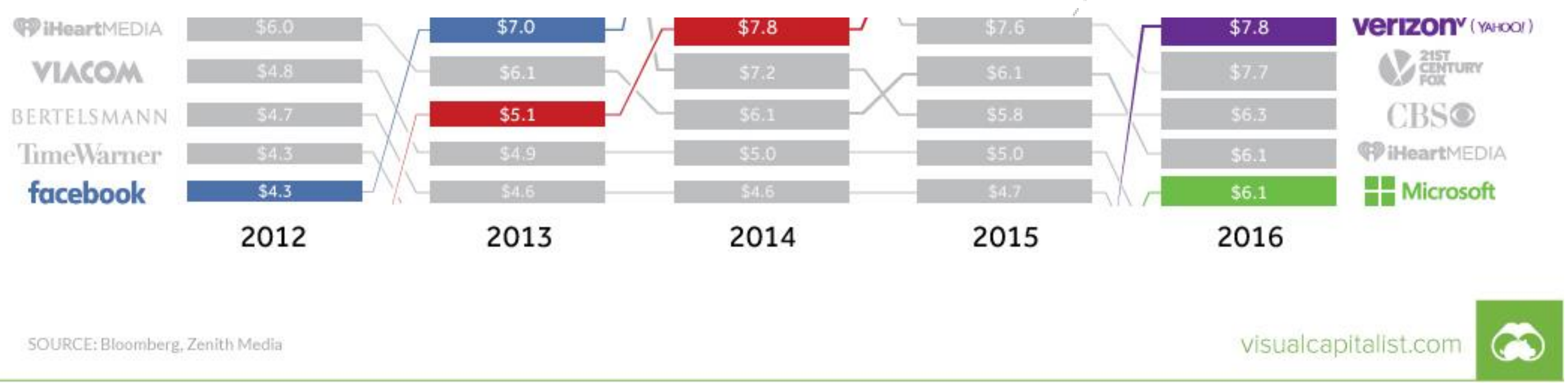
Online Advertising

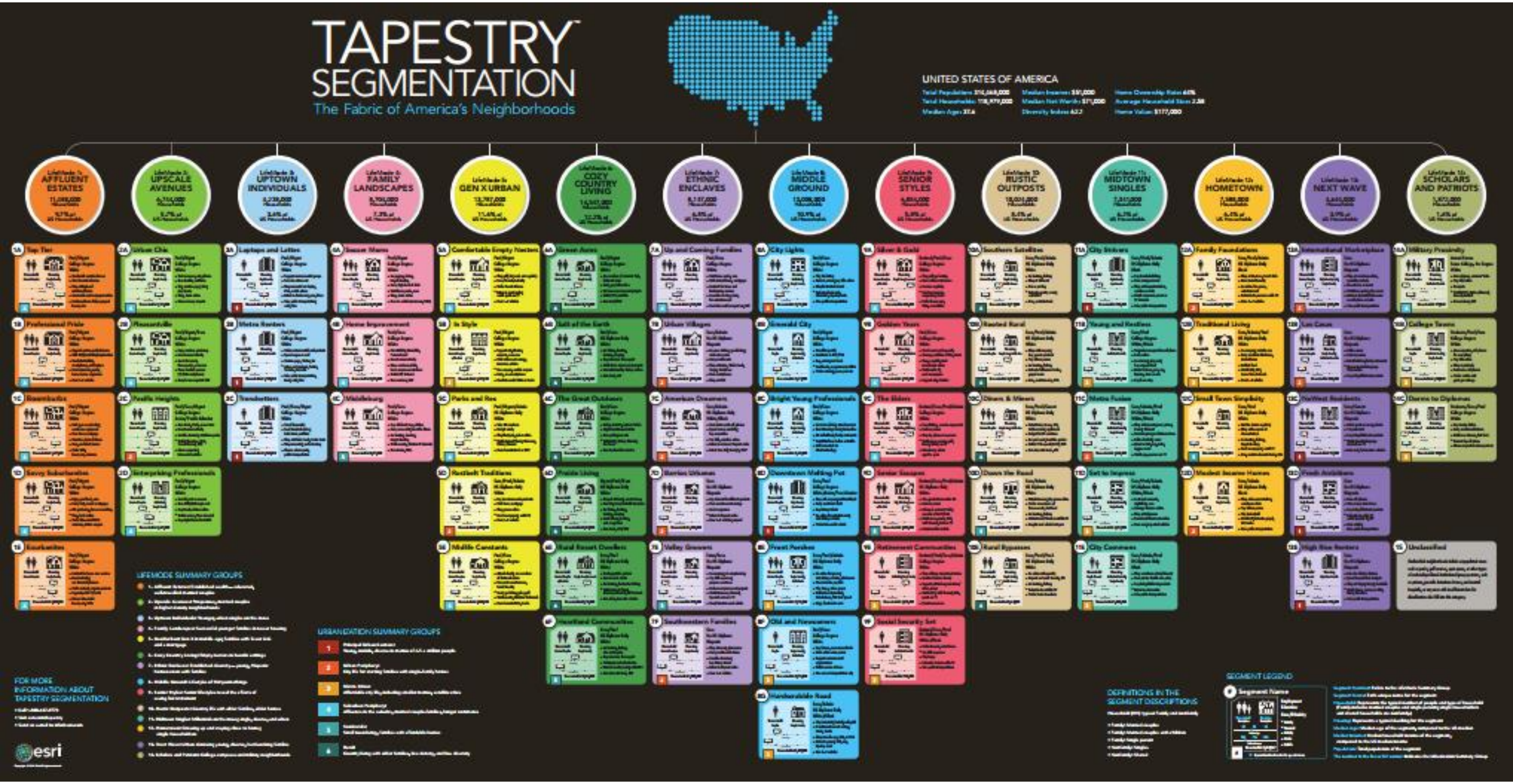


Online Advertising

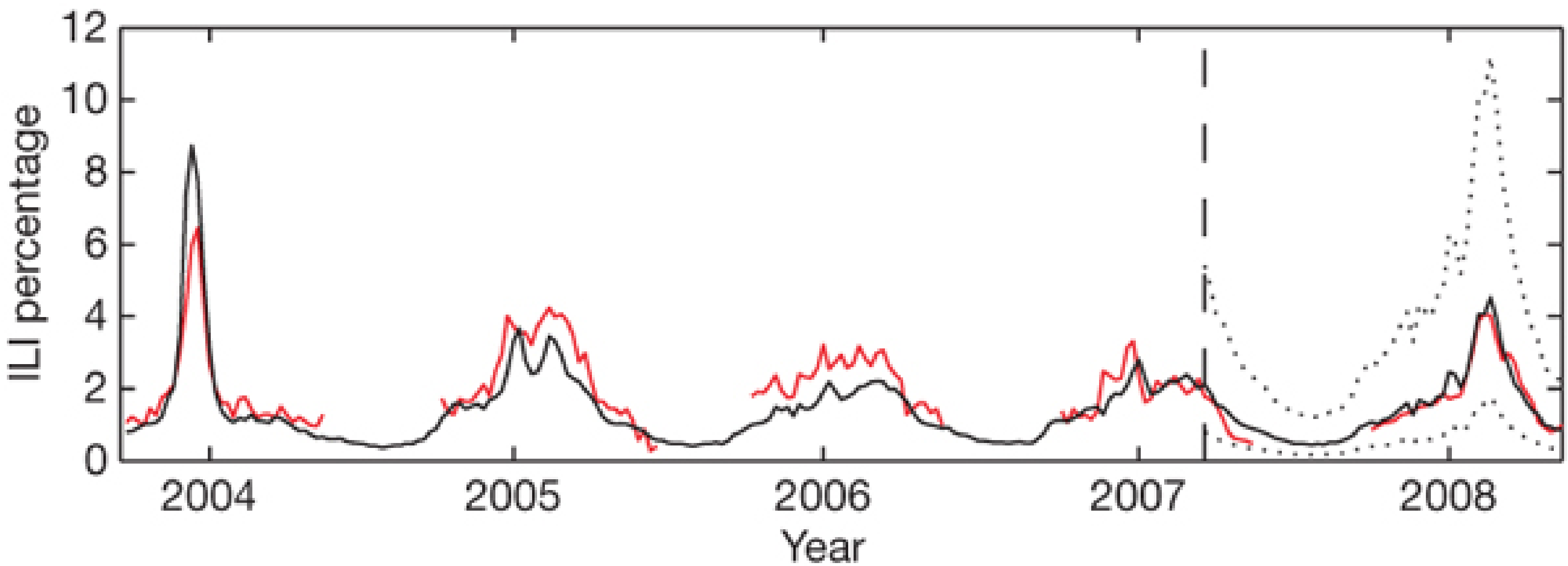
Ad-Supported Internet Brings Over $1 Trillion To The U.S. Economy, Representing 6 Percent Of Country's Total GDP, According To IAB Study Led By Harvard Business School Professor

03.15.17




TAPESTRY SEGMENTATION
The Fabric of America's Neighborhoods
UNITED STATES OF AMERICA
AFFLUENT ESTATES
UPSCALE AVENUES
UPTOWN INDIVIDUALS
FAMILY LANDSCAPES
GEN X URBAN
COZY COUNTRY LIVING
ETHNIC ENCLAVES
MIDDLE GROUND
SENIOR STYLES
RUSTIC OUTPOSTS
MIDTOWN SINGLES
HOMETOWN
NEXT WAVE
SCHOLARS AND PATRIOTS
LIFEMODE SUMMARY GROUPS
URBANIZATION SUMMARY GROUPS
SEGMENT LEGEND
esri

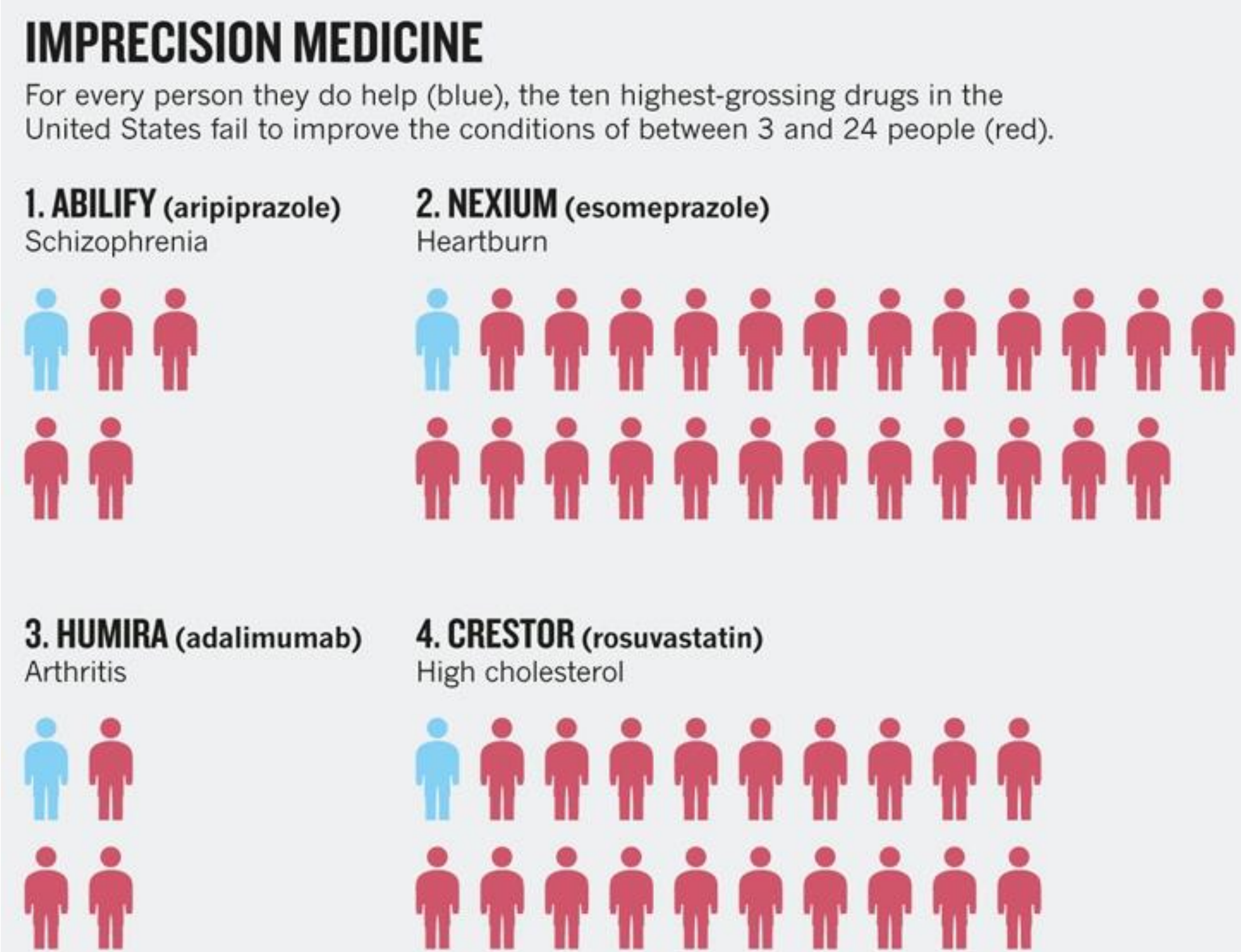
Health



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

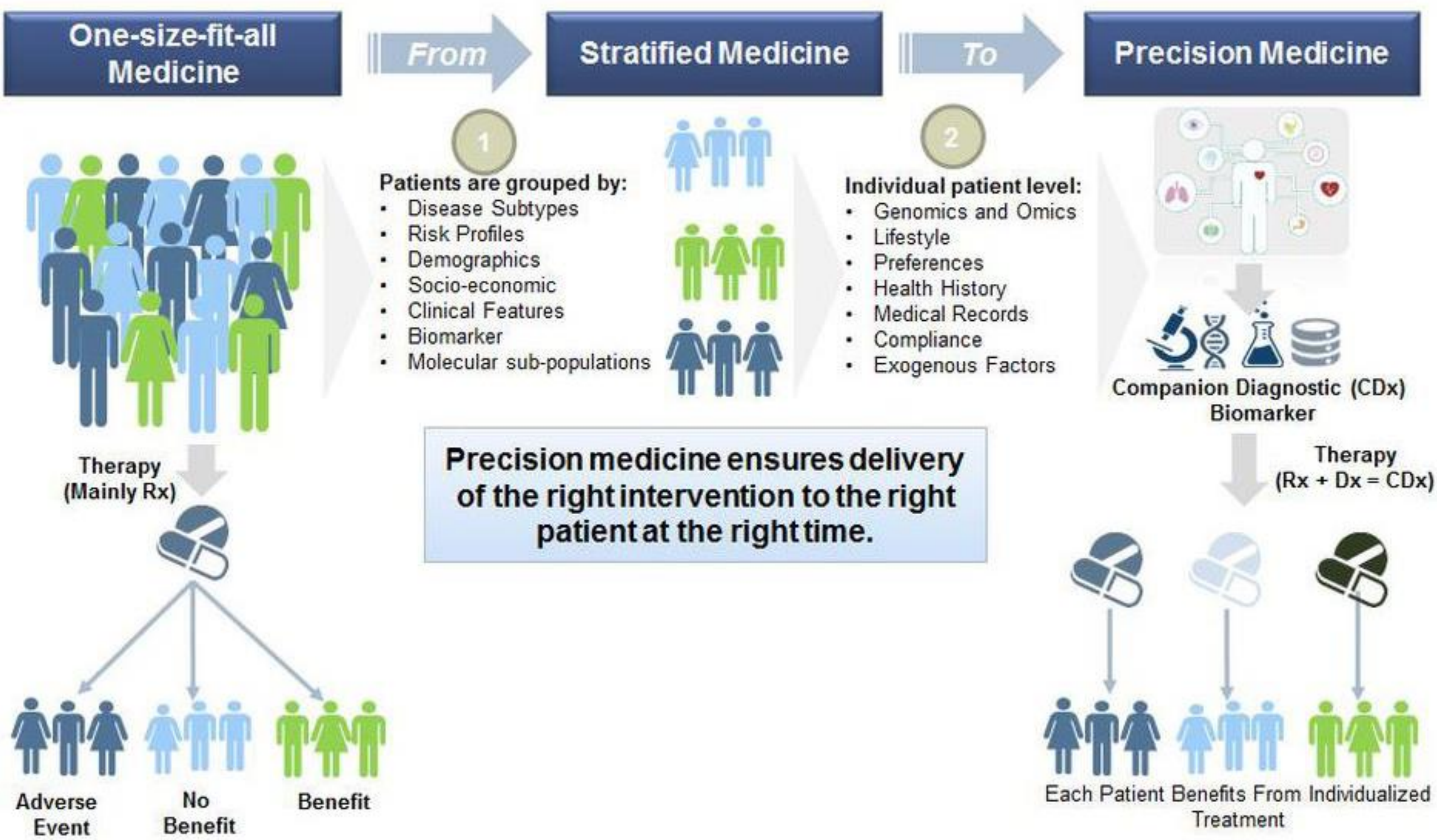
Detecting influenza epidemics using search engine query data

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



https://**www.nature.com**/news/personalized-medicine-time-for-one-person-trials-1.17411

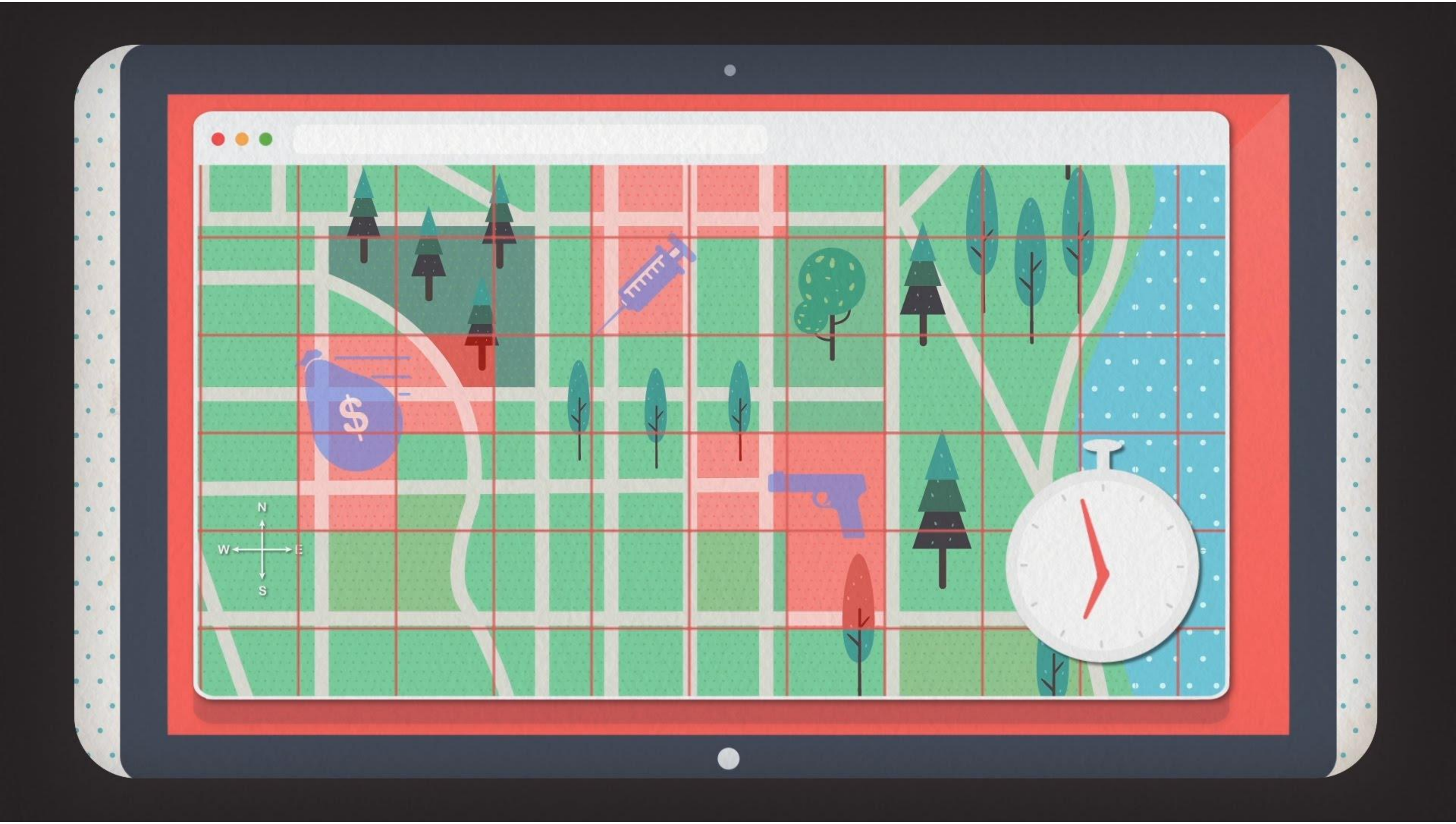
Precision Medicine



Source: forbes.com

Predictive Policing

Predictive Policing



The dark side of the force…

http://ragekg.deviantart.com/art/The-Dark-Side-of-the-Force-

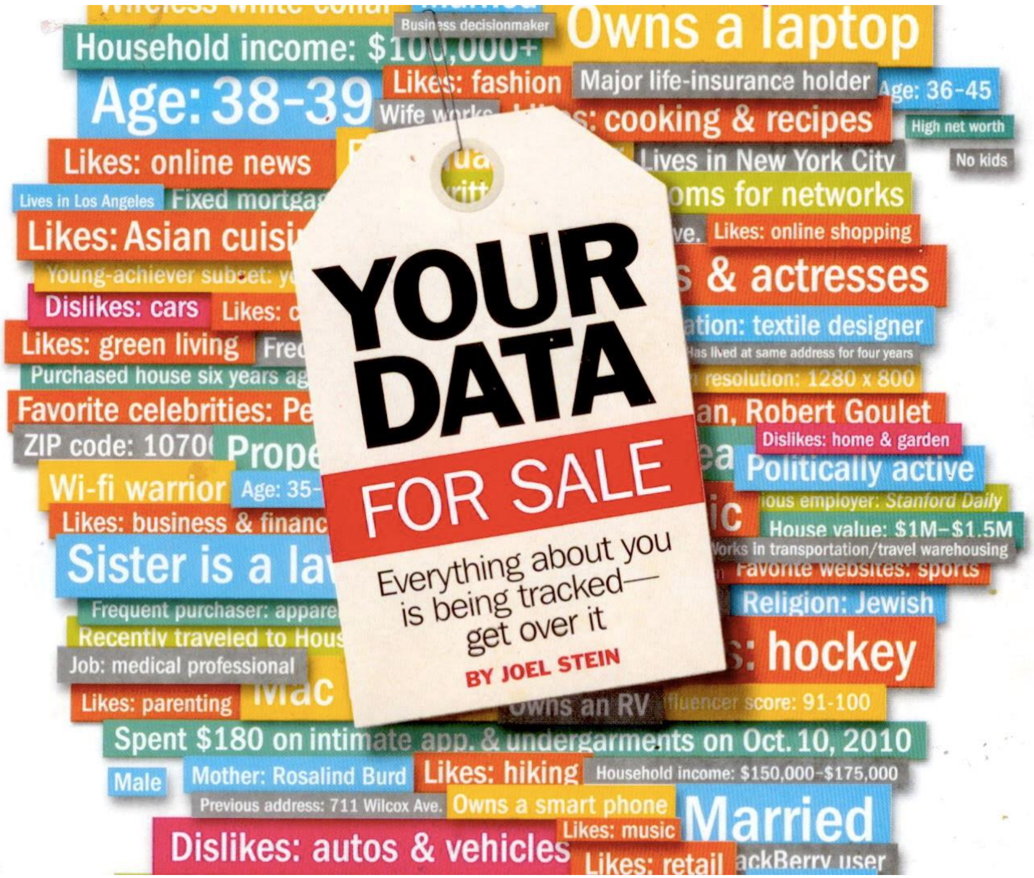
39% of the experts agree…

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

- Medical records
- Insurance records

- Search logs
- Browse logs
- Shopping histories

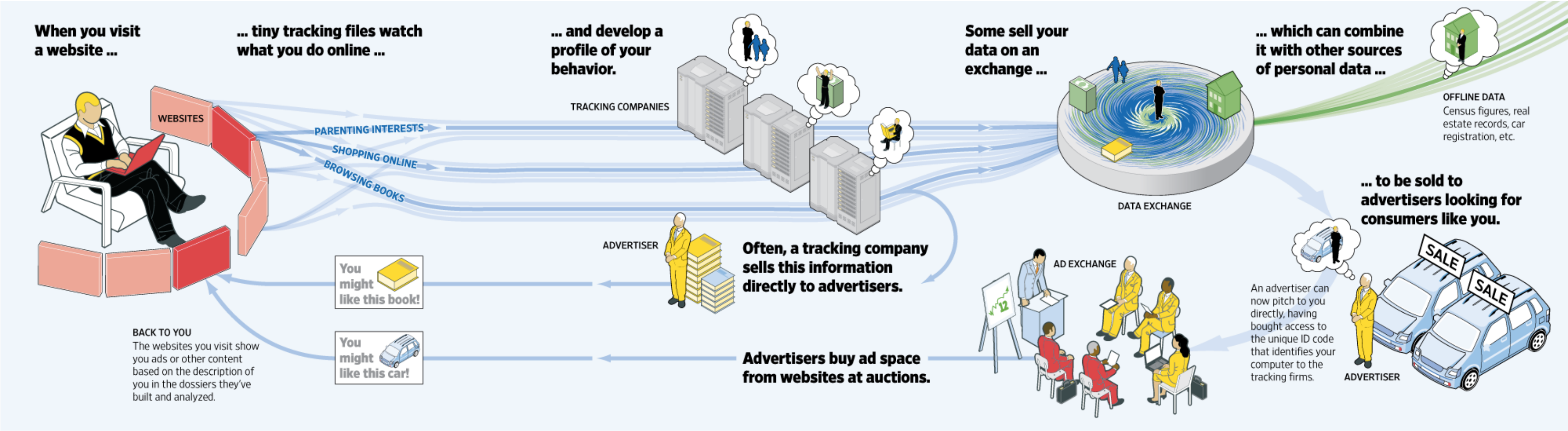
- Photos
- Videos

- Smart phone Sensors
- Mobility trajectories

- …

Very sensitive information …

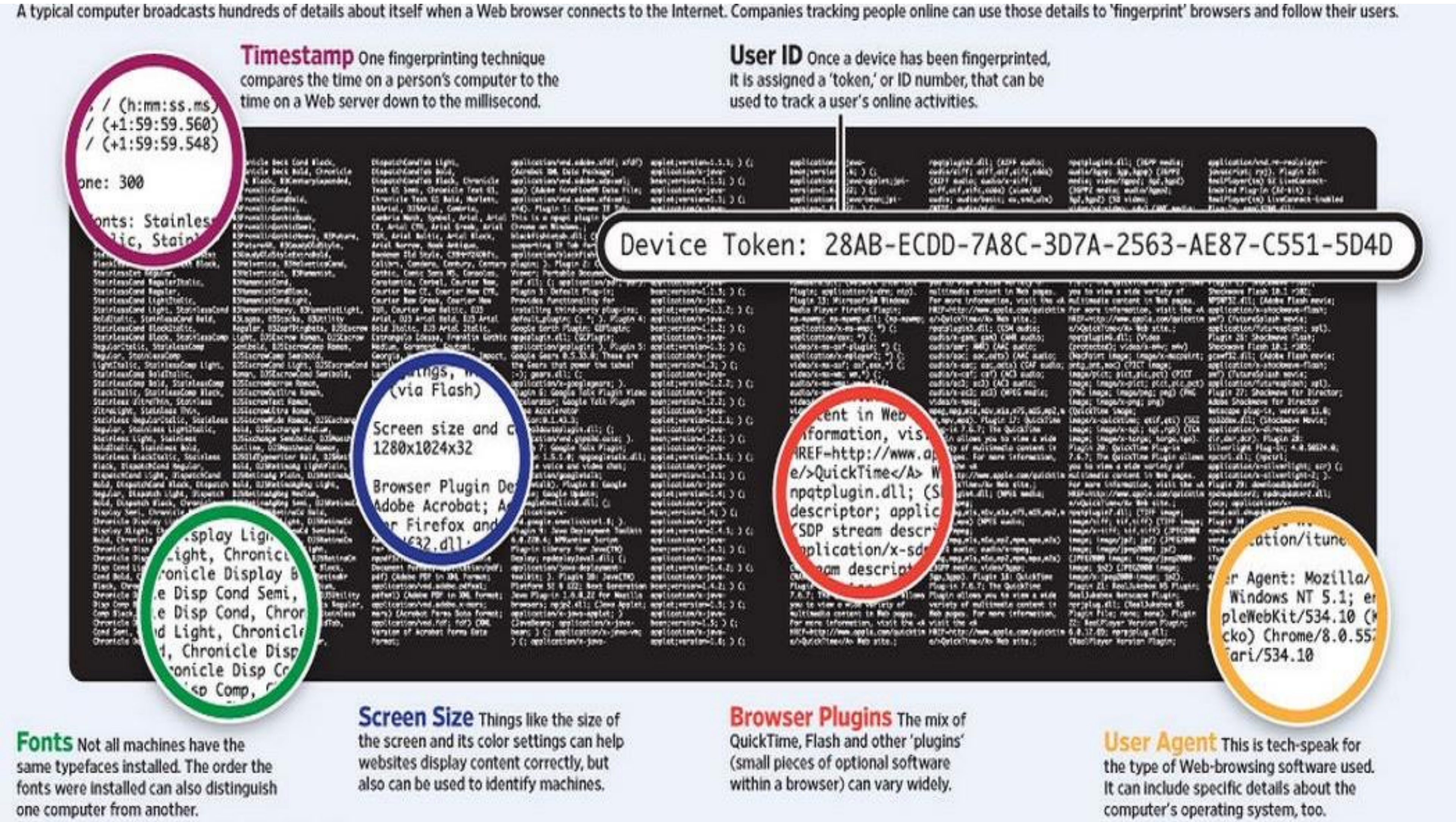
How is this data collected?



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

Isn’t my data anonymous ?

Device Fingerprinting



PANOPTICLICK 3.0

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://panopticlick.eff.org/

Let’s get rid of unique identifiers …

HIPAA
COMPLIANT

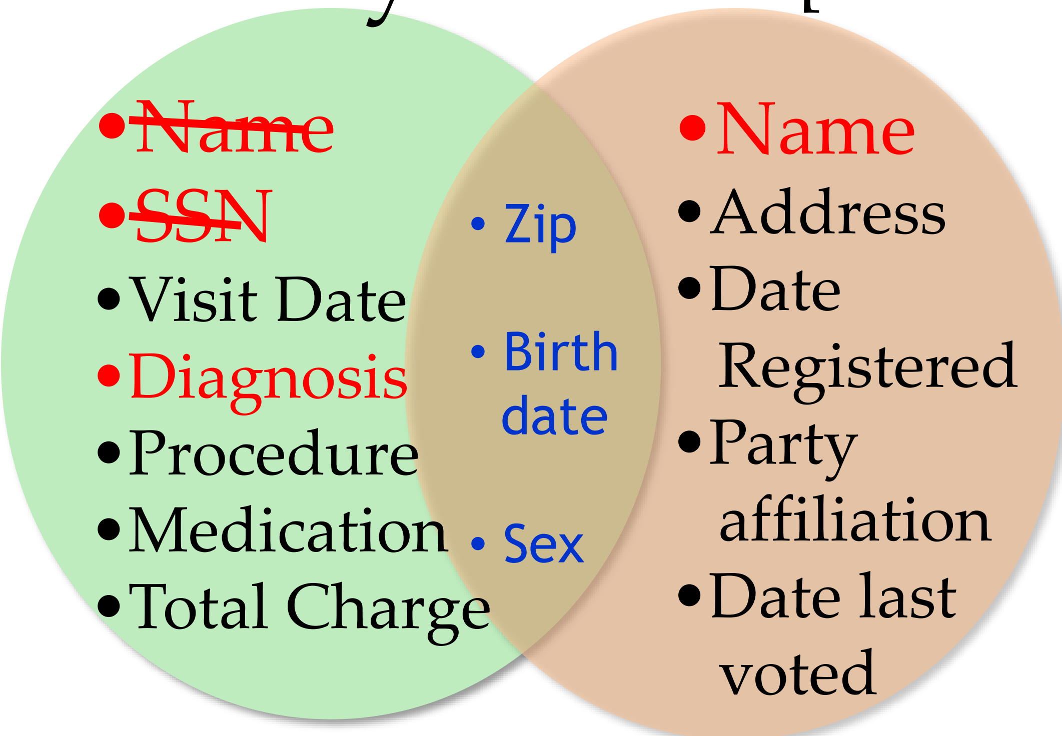
The Massachusetts Governor Privacy Breach [Sweeney IJUFKS 2002]

- ~~Name~~
- ~~SSN~~
- Visit Date
- Diagnosis
- Procedure
- Medication
- Total Charge

- Zip
- Birth date
- Sex

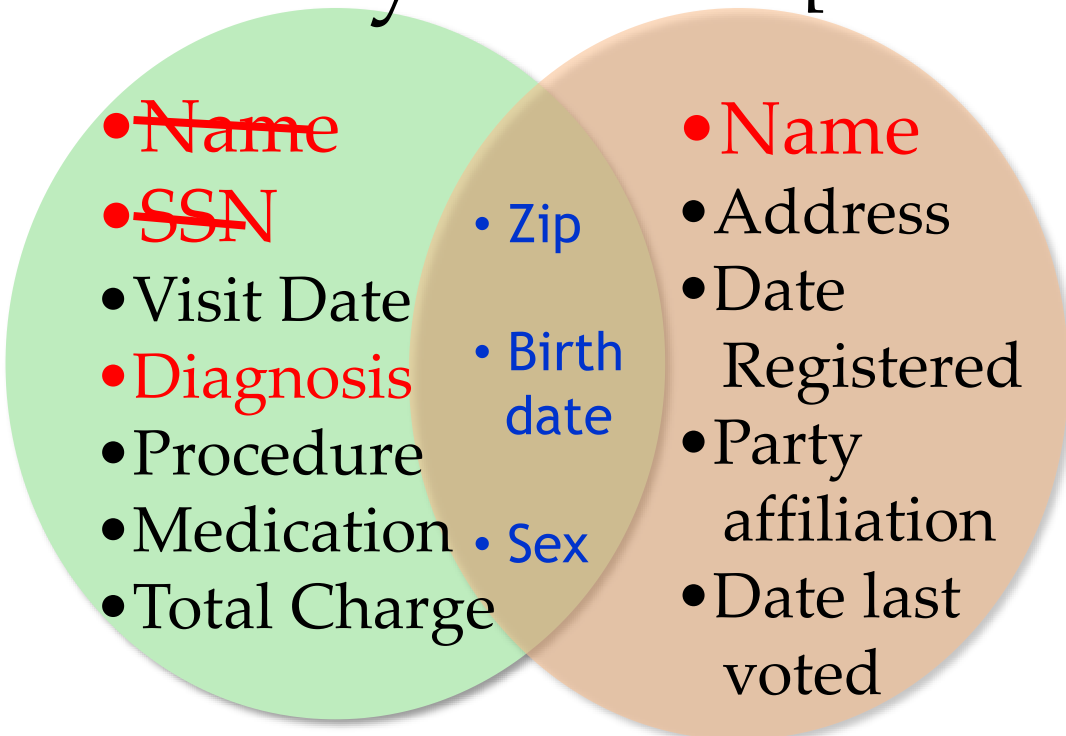
Medical Data

The Massachusetts Governor Privacy Breach [Sweeney IJUFKS 2002]



Medical Data **Voter List**

The Massachusetts Governor Privacy Breach [Sweeney IJUFKS 2002]

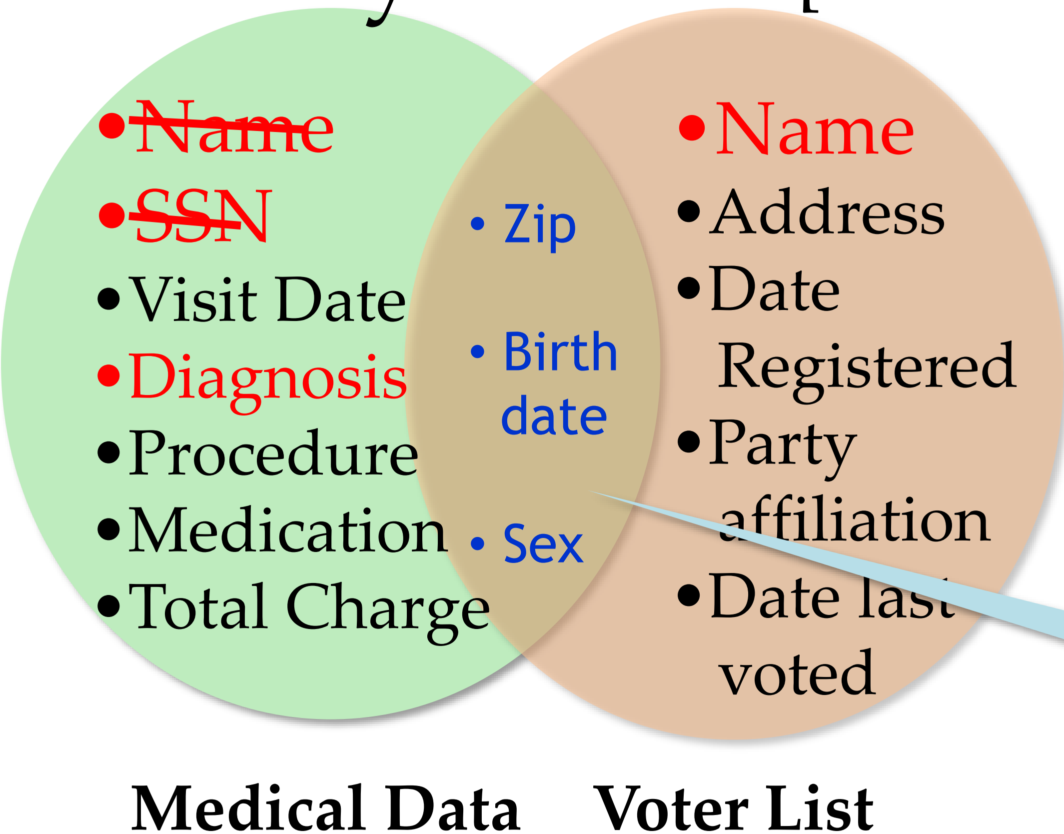


Medical Data **Voter List**

- 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

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 deFinitti's theorem
Jane12345	Zombie games
Jane12345	Warcraft
Jane12345	Beatles anthology
Jane12345	Ubuntu breeze
Bob222	Python in thought
Bob222	Enthought 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 deFinitti's theorem
112765410	Zombie games
112765410	Warcraft
112765410	Beatles anthology
112765410	Ubuntu breeze
865712345	Python in thought
865712345	Enthought 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

+

Online Data

- 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**

	Number of Impressions
+ Who are interested in **Men**	25
+ Who are interested in **Women**	0

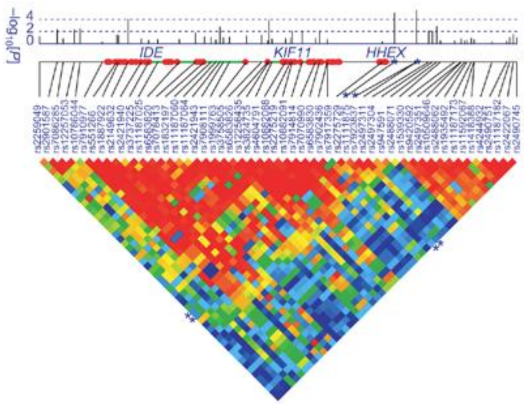
Facebook's learning algorithm uses private information to predict match to ad

[Korolova JPC 2011]

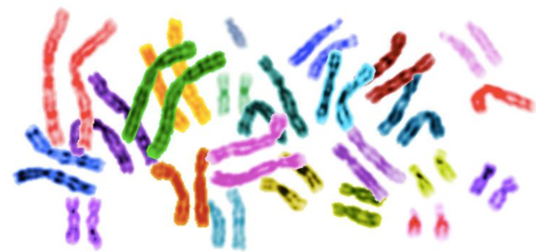
Genome wide association studies

[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

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?

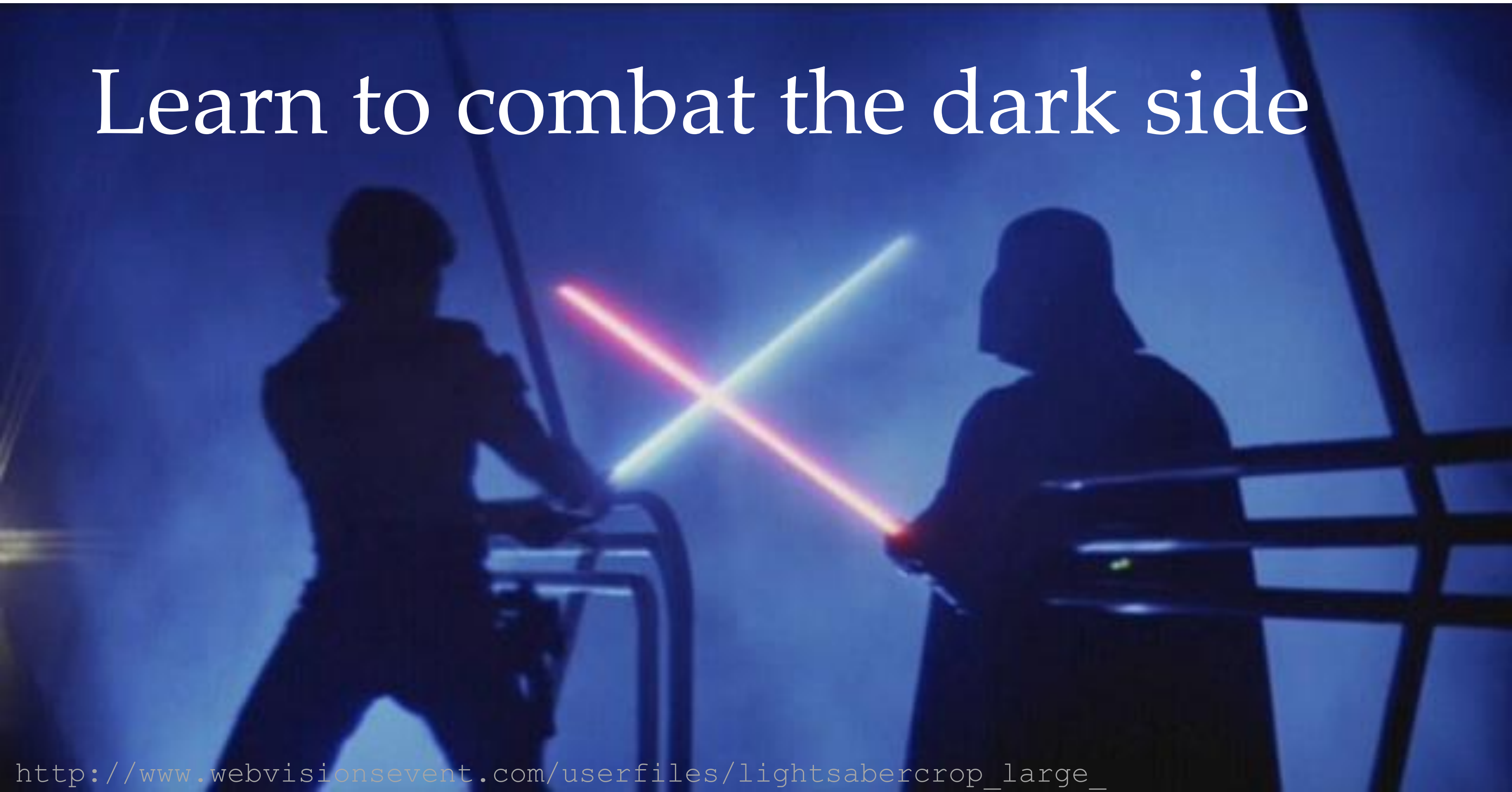
We can with **shadow training**:

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.

Dataset	*Training Accuracy*	*Testing Accuracy*	*Attack Precision*
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
http://www.webvisionsevent.com/userfiles/lightsabercrop_large_

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
  - 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

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

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

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

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

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

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/
- Paper critiques are individual work and submission.
- All suspected cases of violation will be aggressively pursued