

SHOULD WE HIDE FROM BIG DATA?

Westmont Downtown
April 13, 2017

Prof. Donald J. Patterson
Department of Math and
Computer Science

A decorative graphic in the bottom right corner of the slide. It features a stylized film reel with circular perforations and a film strip with a grid pattern, both rendered in a light blue color that blends with the background.

INTRODUCTIONS

PROFESSIONAL BACKGROUND

- B.S., Computer Science from Cornell (NY)
- M.Eng, Electrical Engineering
- Tomahawk STRIKE Officer
 - USS CURTIS WILBUR, DDG-54, Japan
- Operations Officer/Navigator
 - USS SIMON LAKE, AS-33, Italy
- Ph.D in Computer Science and Engineering at UW
- Director of the Laboratory for Ubiquitous Computing and Interaction
- Co-founder of quub, whisper.fm, swayr, waitscout, audia



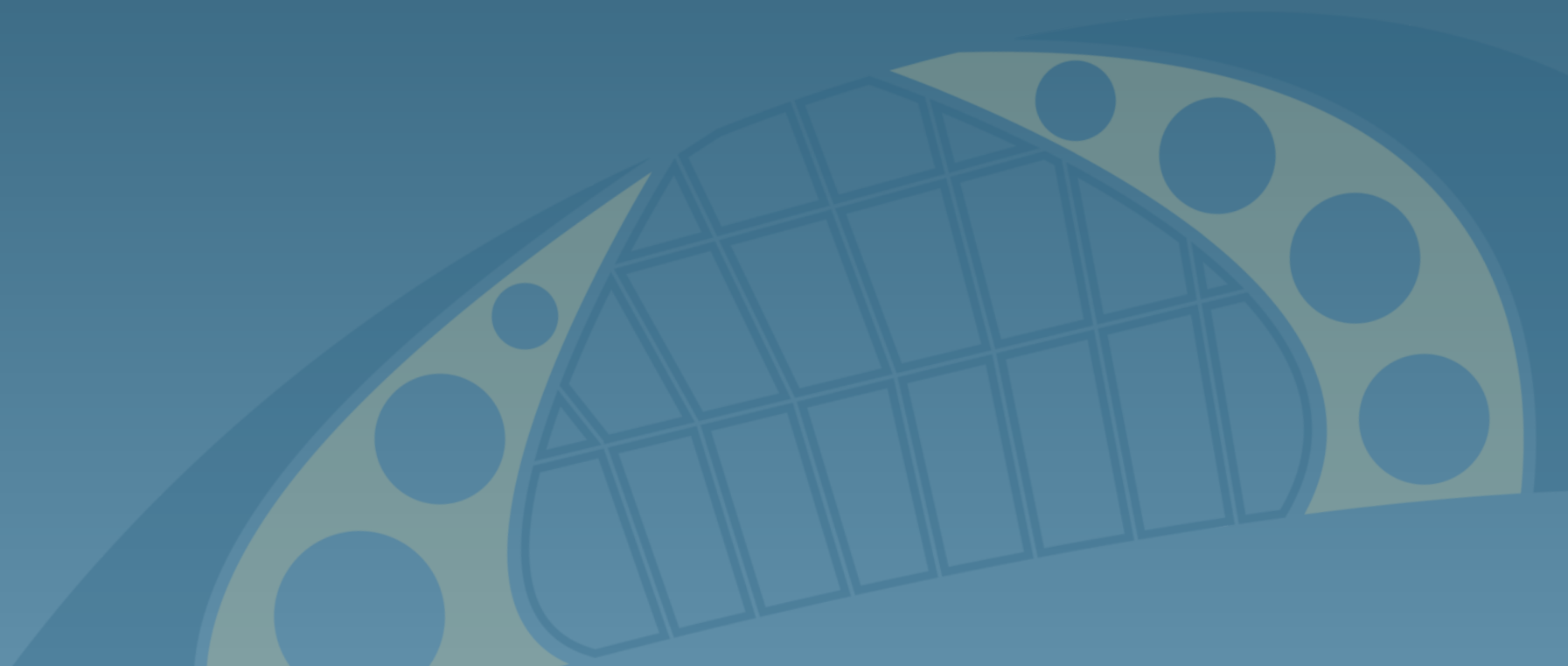
INTRODUCTIONS

WESTMONT

- Joined in 2015 - Launched WICL



BIG DATA



BIG DATA

- WHAT IS BIG DATA?
- HOW IS BIG DATA USED?
- THE GOOD, THE BAD, THE UGLY
- TWO WAYS TO THINK OF BIG DATA
 - PANOPTICON
 - CAPTURE MODEL
- HOW THEN OUGHT WE TO LIVE?

WHAT IS BIG DATA?

Big data is a broad term for data sets so large or complex that traditional data processing applications are inadequate.

There is so much that it is hard to capture, store, curate, analyze, search, share, transfer, visualize, and keep it all private.

WHAT IS BIG DATA?

Big data is also a business term that means using lots of data to do

smarter

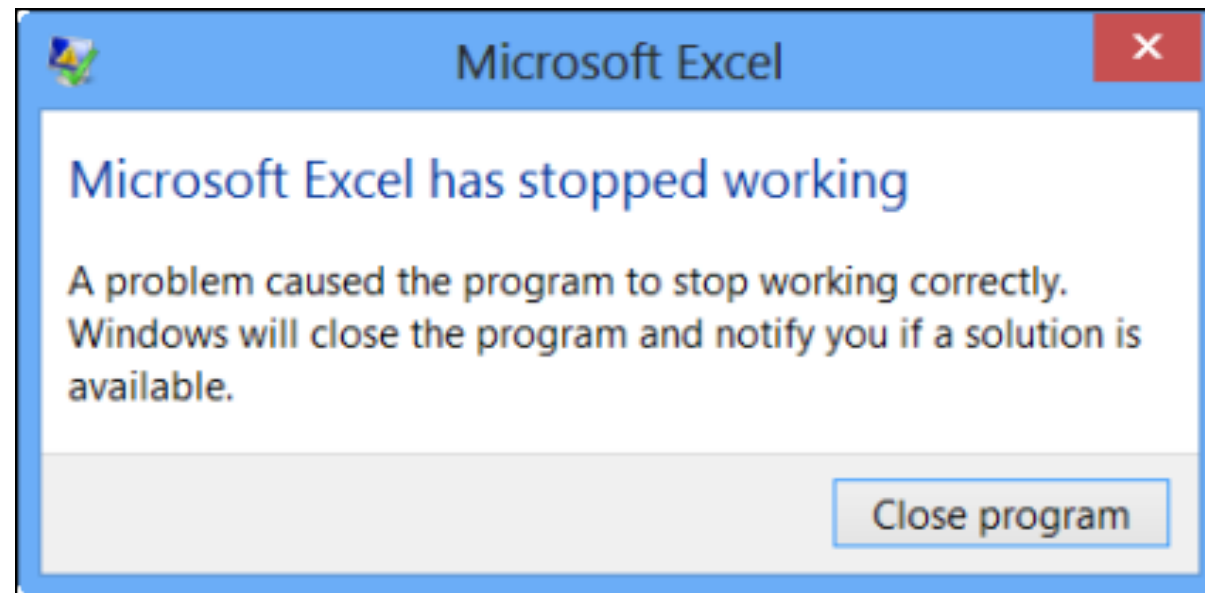
faster

more insightful

data analysis

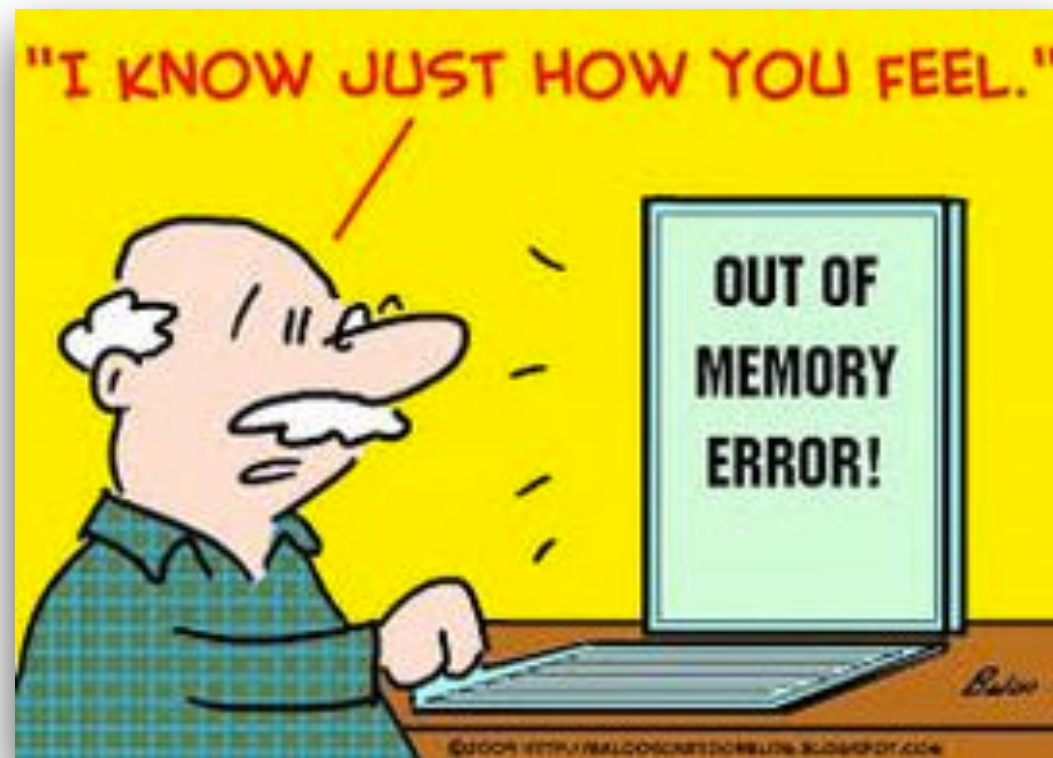
SYMPTOMS OF BIG DATA

- Microsoft Excel crashes



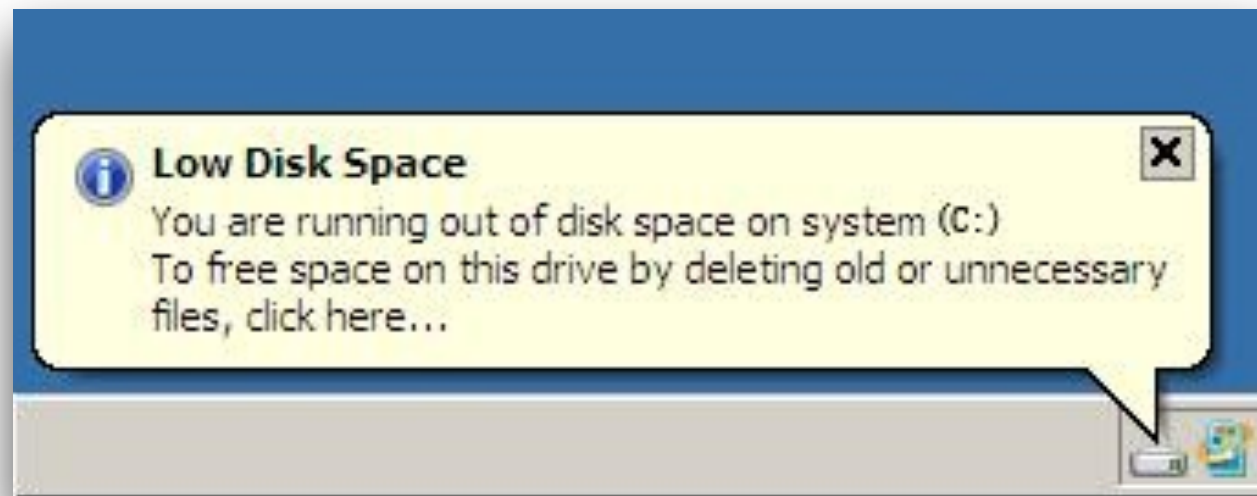
SYMPTOMS OF BIG DATA

- Microsoft Excel crashes
- The data cannot fit in the **memory** of one computer



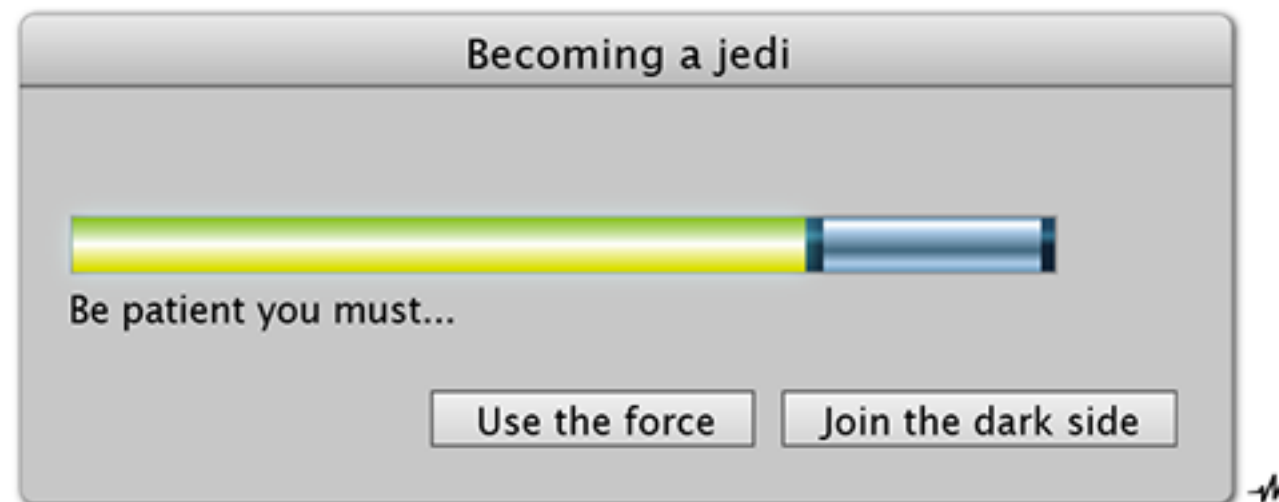
SYMPTOMS OF BIG DATA

- Microsoft Excel crashes
- The data cannot fit in the **memory** of one computer
- The data cannot fit on the **disk** of one machine



SYMPTOMS OF BIG DATA

- Microsoft Excel crashes
- The data cannot fit in the **memory** of one computer
- The data cannot fit on the **disk** of one machine
- It would take several **days, months,** or **years** to process the data

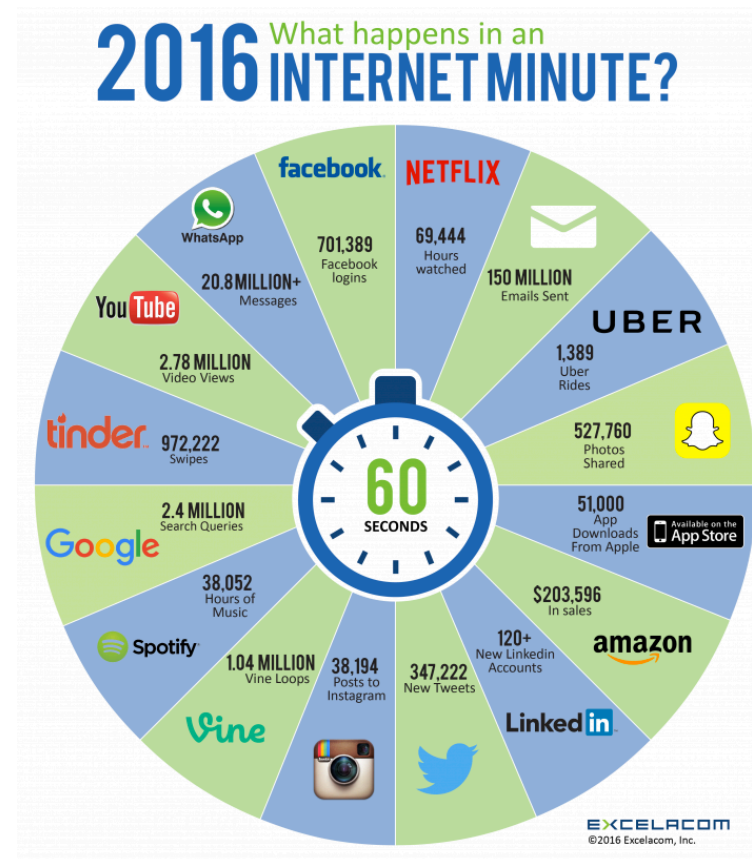


SYMPTOMS OF BIG DATA

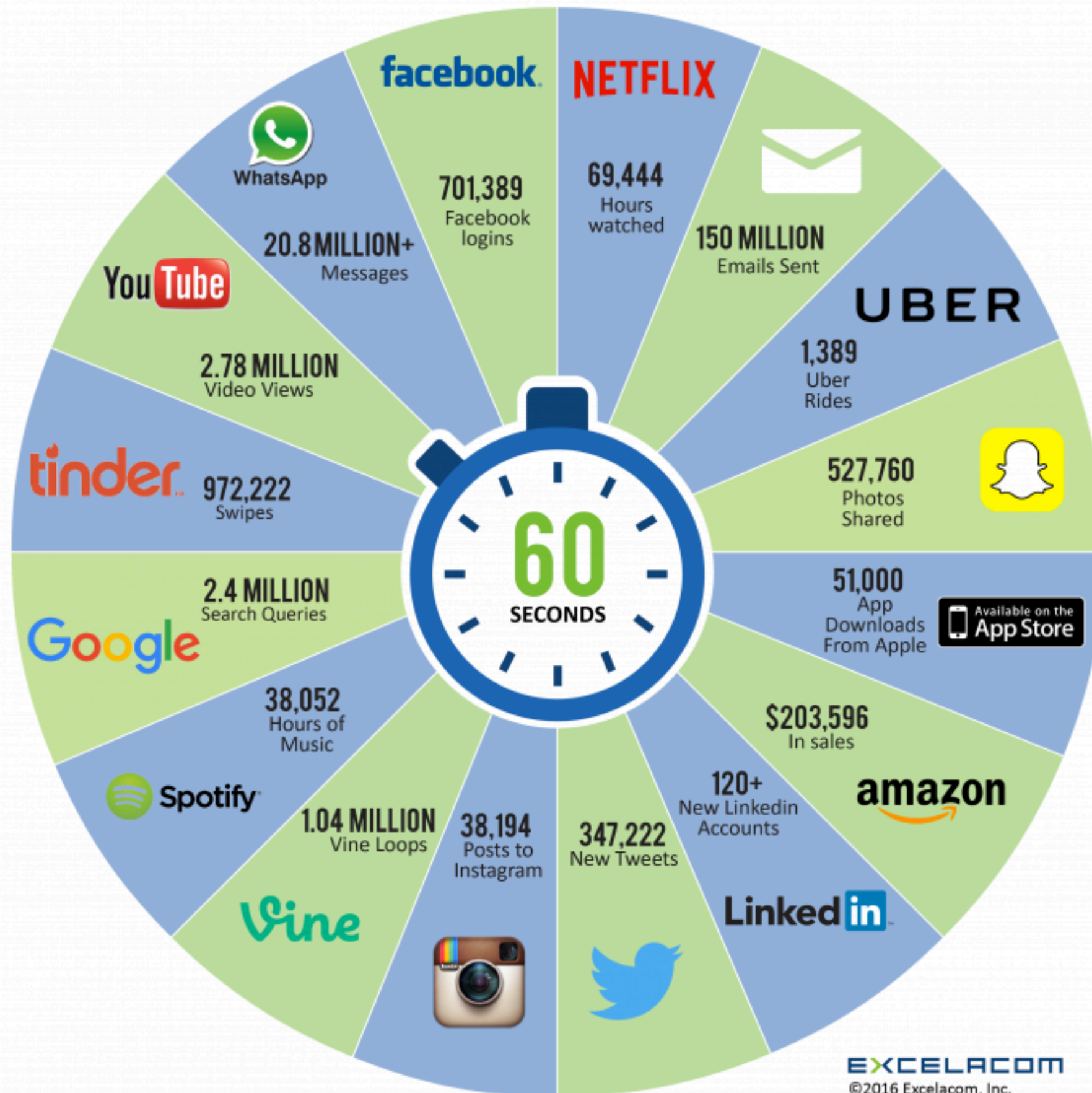
- Microsoft Excel crashes
- The data cannot fit in the **memory** of one computer
- The data cannot fit on the **disk** of one machine
- It would take several **days, months,** or **years** to process the data
- You are using words like terabytes, petabytes, exabytes to describe your data

1 exabyte is 15,112,000,000,000,000 bytes

WHERE DOES BIG DATA COME FROM?

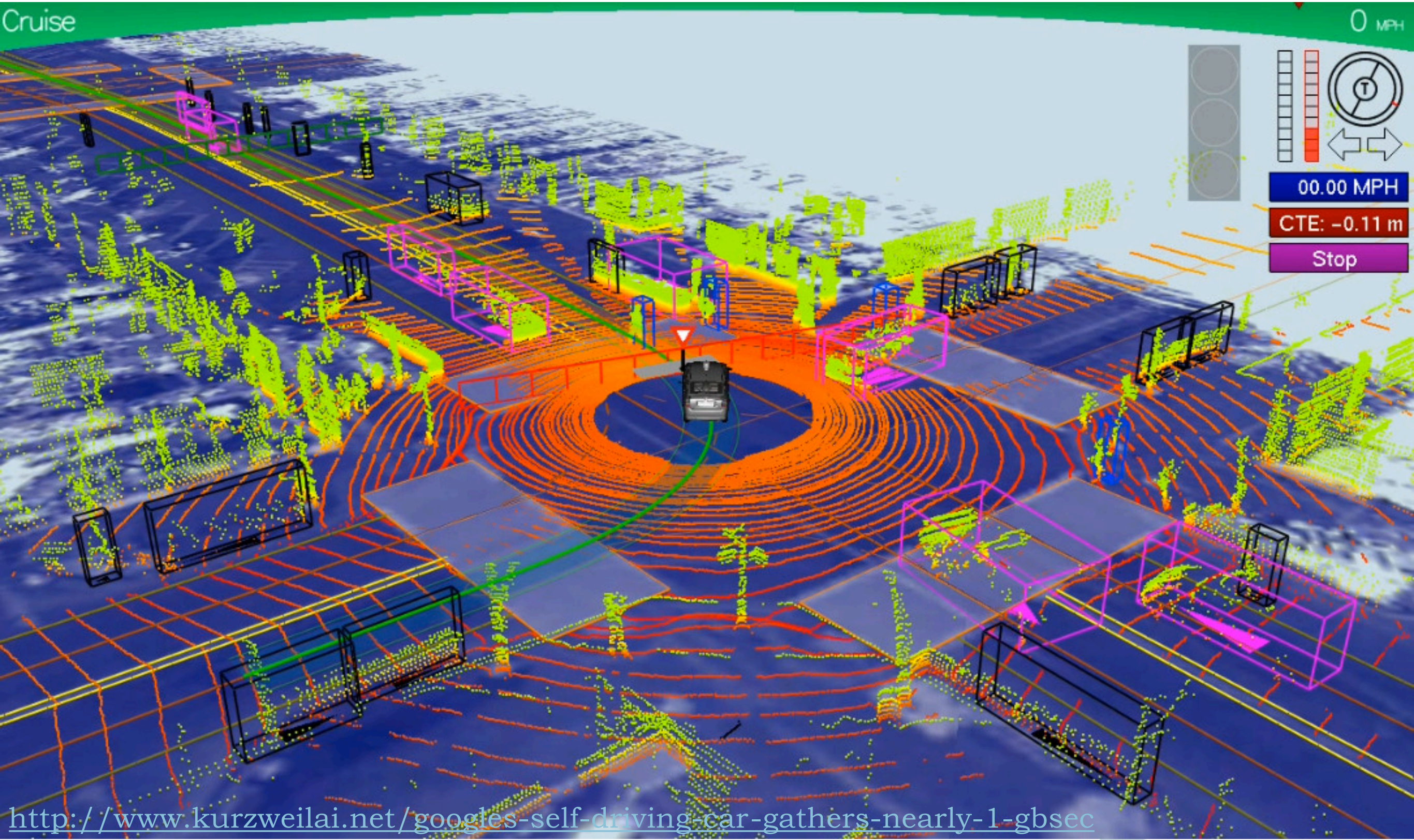


2016 What happens in an INTERNET MINUTE?



WHERE DOES BIG DATA COME FROM?

Google self-driving cars create 45 GB per minute



BIG DATA

- WHAT IS BIG DATA?
- HOW IS BIG DATA USED?
- THE GOOD, THE BAD, THE UGLY
- TWO WAYS TO THINK OF BIG DATA
 - PANOPTICON
 - CAPTURE MODEL
- HOW THEN OUGHT WE TO LIVE?

WHAT HAPPENS WITH THAT DATA

- Data Generation
 - users are browsing, buying and reviewing products

The screenshot displays the Amazon website interface for a Weber grill. At the top, the Amazon Prime logo is on the left, and the search bar contains "weber grill". To the right of the search bar, it says "amazon@IslaVista" and "Convenient Order P". Below the search bar, there are links for "Departments", "Amazon@IslaVista", and "Browsing History". On the right side of the top navigation bar, there are links for "EN", "Hello, Donald", "Account & Lists", "Orders", "Prime", and a shopping cart icon with "0" items.

The main content area features a banner for "FREE ASSEMBLY only on select grills" with three grill images and a "Shop now" button. Below this, a link says "Back to search results for 'weber grill'".

The product page for the "Weber 14401001 Original Kettle Premium Charcoal Grill, 22-Inch, Black" is shown. It includes a large image of the grill, a smaller image strip on the left, and a "Roll over image to zoom in" text. The product title is "Weber 14401001 Original Kettle Premium Charcoal Grill, 22-Inch, Black". It has a 4.5-star rating from 743 customer reviews and 83 answered questions. The price is \$149.00, with a Prime logo. A promotional message states: "Get \$50.00 off instantly: Pay \$99.00 upon approval for the Amazon Rewards Visa Card. Learn more". The status is "In Stock." and it says "Want it Friday, April 14? Order within 12 hrs 53 mins and choose Standard Shipping at checkout. Details". It also mentions "Ships from and sold by Amazon.com. Gift-wrap available."

Below the product image, there are color options: "Color: Black" with three small grill images (black, red, and silver). The "Style" is listed as "Grill".

On the right side of the product page, there are social sharing links (Share, Email, Facebook, Twitter, Pinterest). Below these, there is a quantity selector set to "1" and an "Add to Cart" button. A link says "Turn on 1-Click ordering for this brows". The shipping information shows "Ship to: Sandra S Patterso- Santa Barba - 93108". There are buttons for "Add to List" and "Add to your Dash Buttons". At the bottom right, there is a section for "Other Sellers on Amazon" showing the price "\$149.00" and an "Add to C" button.

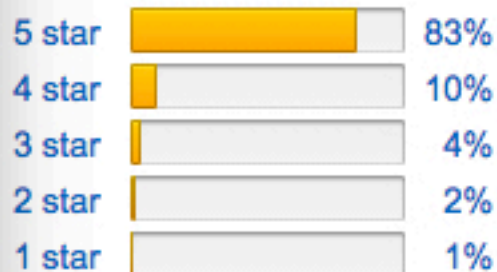
WHAT HAPPENS WITH THAT DATA

- Data Generation
 - users are browsing, buying and reviewing products
- Data Uses

Customer Reviews

★★★★★ 743

4.7 out of 5 stars ▾



Share your thoughts with other customers

Write a customer review

[See all verified purchase reviews ▸](#)

Top Customer Reviews

★★★★★ **First time griller success!**

By [DesignGirl28](#) on May 25, 2015

Color: Black | Style: Grill | **Verified Purchase**

First off, I have to tell you that this is the first grill I have ever used. As a single woman, I was never interested in grilling until I bought a house two years ago and built a deck in the backyard. With all that outdoor space to enjoy, I just knew I needed a grill for entertaining (and feeding) family and friends. I thought I might buy a grill and just let my dad or brother handle the grilling, but much to my surprise I really got into it and loved grilling hamburgers and hot dogs on this Weber kettle grill for my first backyard party just yesterday!

Customer Images



[See all customer images](#)

Most Recent Customer Reviews

★★★★☆ **Hoping its Build Quality is Better than the Quality of the Instructions**

My wife and I just assembled this grill. The included instructions are a joke. It is a booklet containing drawings and a couple of smiley faces. [Read more](#)

Published 3 days ago by A. Petrone

★★★★★ **Fantastic product and service**

Fantastic product and service! The grill arrived via FedEx earlier than predicted when I ordered it. I used the grill almost every night and love it!

Published 6 days ago by Julia

WHAT HAPPENS WITH THAT DATA

- Data Generation
 - users are browsing, buying and reviewing products
- Data Uses

Customers who viewed this item also viewed

Page



Weber 14501001 Master-Touch Charcoal Grill, 22-Inch, Black

★★★★★ 100

\$199.00 ✓Prime



Weber 741001 Original Kettle 22-Inch Charcoal Grill

★★★★★ 768

\$72.00



Weber 16401001 Original Kettle Premium Charcoal Grill, 26-Inch, Black

★★★★★ 115

\$299.00 ✓Prime



Weber 15301001 Performer Charcoal Grill, 22-Inch, Black

★★★★★ 83

\$249.00 ✓Prime



Weber 15501001 Performer Deluxe Charcoal Grill, 22-Inch, Black

★★★★★ 291

\$399.00 ✓Prime



WHAT HAPPENS WITH THAT DATA

- Data Generation
 - users are browsing, buying and reviewing products
- Data Uses

Customers who bought this item also bought Page 1



Weber 7149 Grill Cover with Storage Bag for Weber Charcoal Grills, 22.5-Inch
★★★★★ 333
\$29.94 ✓Prime



Weber 7416 Rapidfire Chimney Starter
★★★★★ 4,589
#1 Best Seller in Charcoal Starters
\$14.99 ✓Prime



Classic Accessories 73422 Veranda Kettle-Style Barbecue Grill Cover
★★★★★ 423
\$16.05 ✓Prime



Weber 7451 Standard Kettle Cover, Fits 22-1/2-Inch Charcoal Grills
★★★★☆ 1,030



Weber 7403 Char-Basket Charcoal Briquet Holders
★★★★★ 636
#1 Best Seller in Grill Baskets
\$17.99 ✓Prime

WHAT HAPPENS WITH THAT DATA

- Data Generation
 - users are browsing, buying and reviewing products
- Data Uses

Frequently bought together



Total price: **\$193.93**

Add all three to Cart

Add all three to List

i These items are shipped from and sold by different sellers. [Show details](#)

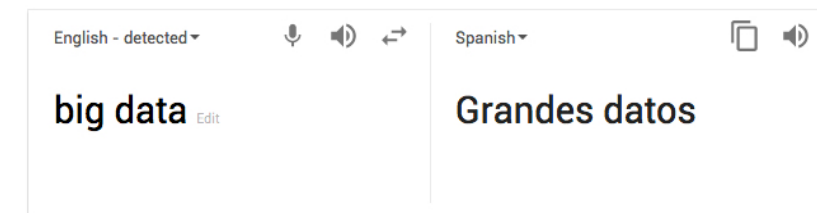
- ✓ **This item:** Weber 14401001 Original Kettle Premium Charcoal Grill, 22-Inch, Black **\$149.00**
- ✓ Weber 7416 Rapidfire Chimney Starter **\$14.99**
- ✓ Weber 7149 Grill Cover with Storage Bag for Weber Charcoal Grills, 22.5-Inch **\$29.94**

BIG DATA

- WHAT IS BIG DATA?
- HOW IS BIG DATA USED?
- THE GOOD, THE BAD, THE UGLY
- TWO WAYS TO THINK OF BIG DATA
 - PANOPTICON
 - CAPTURE MODEL
- HOW THEN OUGHT WE TO LIVE?

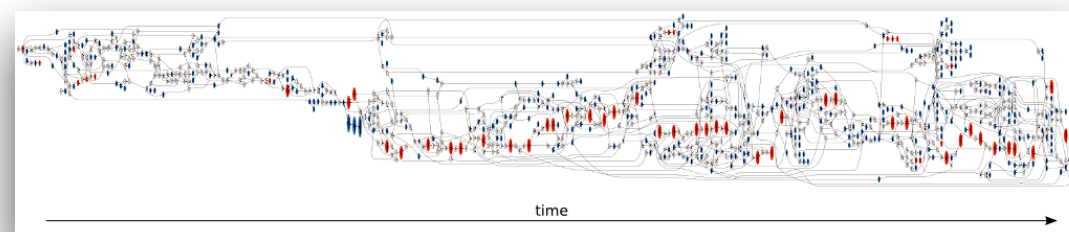
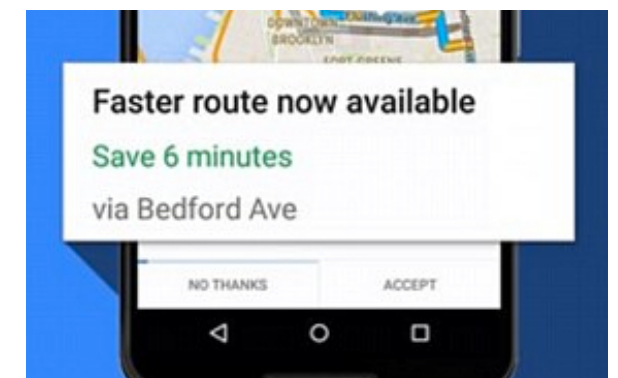
GOOD BIG DATA

- Search engines
- Better inventory control - things are on the shelves
- Automatic translation
- Transformations in genetic and pharmaceutical research
- Traffic rerouting
- Targeted advertising
- “Remind me to get milk on the way home”
- Sex offender bracelets
- Software project management
- Tracking food or hazardous waste



What is GenBank?

GenBank[®] is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences ([Nucleic Acids Research, 2013 Jan;41\(D1\):D36-42](#)). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.





["How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did" Forbes.com](https://www.forbes.com/...)

<https://flic.kr/p/qSqFYA>

BAD BIG DATA

- When sensitive data is discovered without your consent, the potential for coercion multiplies
- Chinese citizen control



DISCONCERTING BIG DATA

- When sensitive data is discovered without your consent, the potential for coercion multiplies
 - Chinese citizen control
 - Revealing alcohol purchases



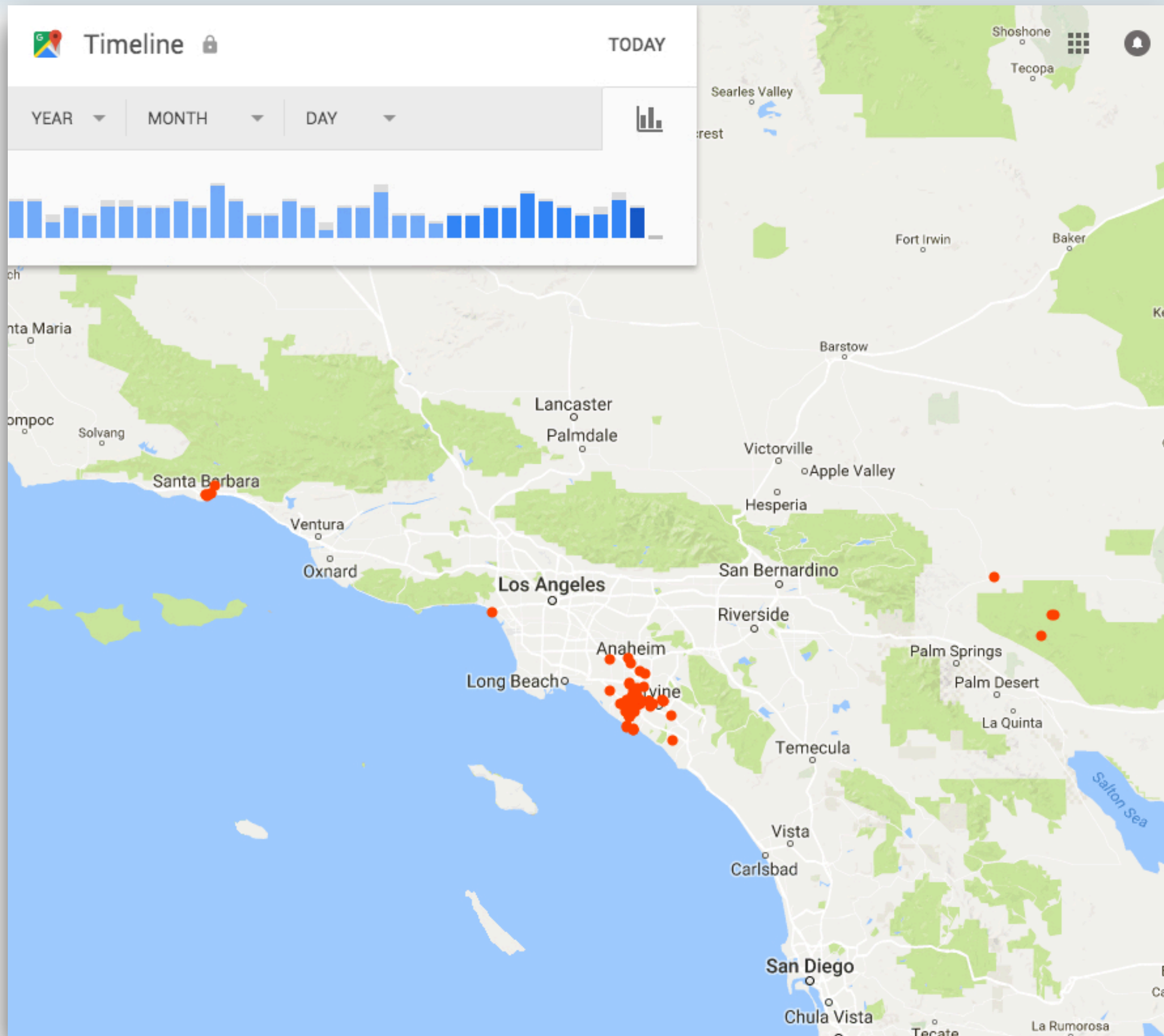
UGLY BIG DATA

IT'S GETTING PHYSICAL

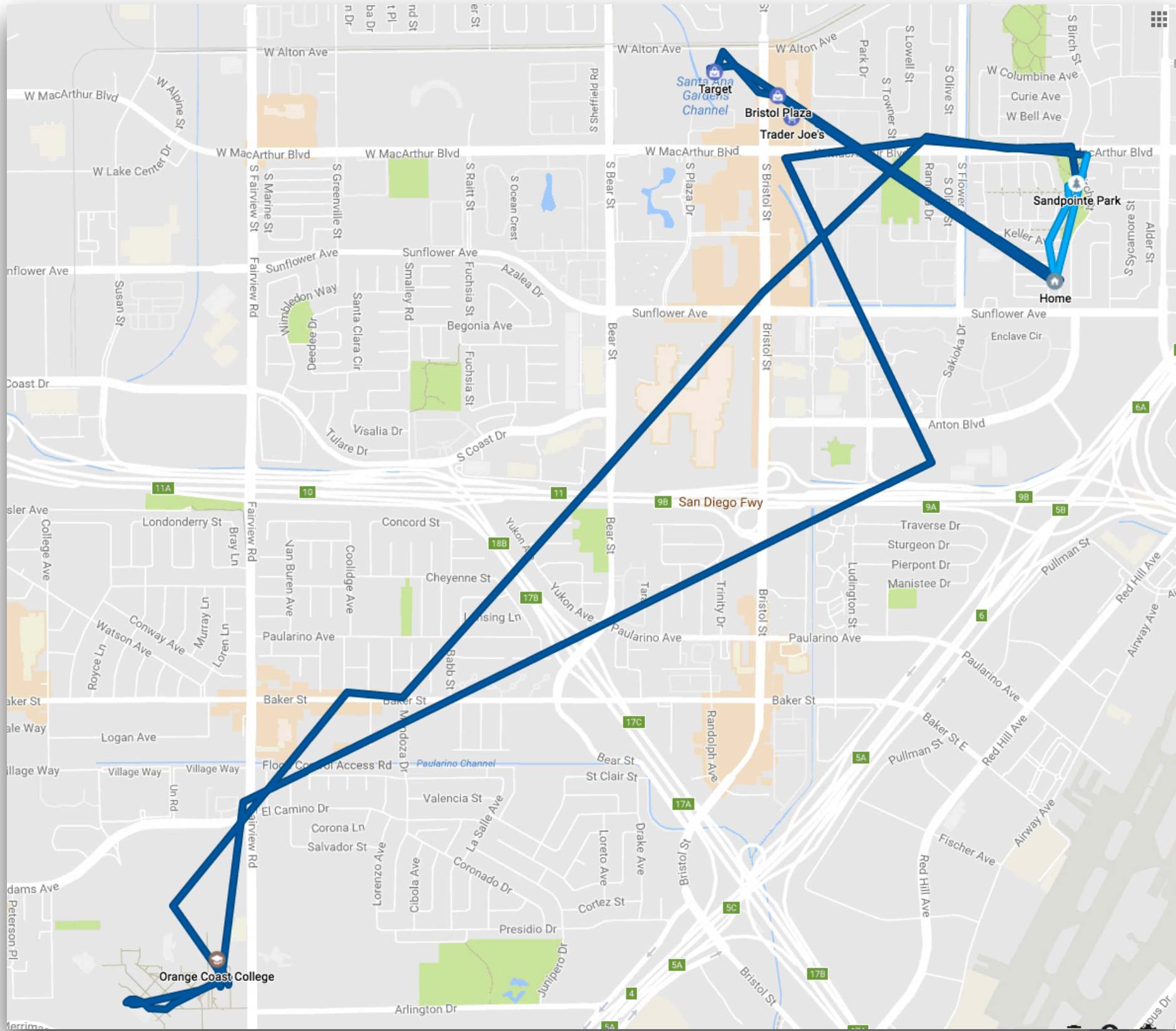
- Increasingly this data is being generated by physical processes not online processes
 - Your license plate is read and automatically analyzed
 - Your ATM transaction anchors you to a location
 - Your phone checks in with a cell tower leaving a pattern of location
 - Your fit bit records location and activity levels



UGLY BIG DATA



UGLY BIG DATA



UGLY BIG DATA



UGLY BIG DATA



UGLY BIG DATA



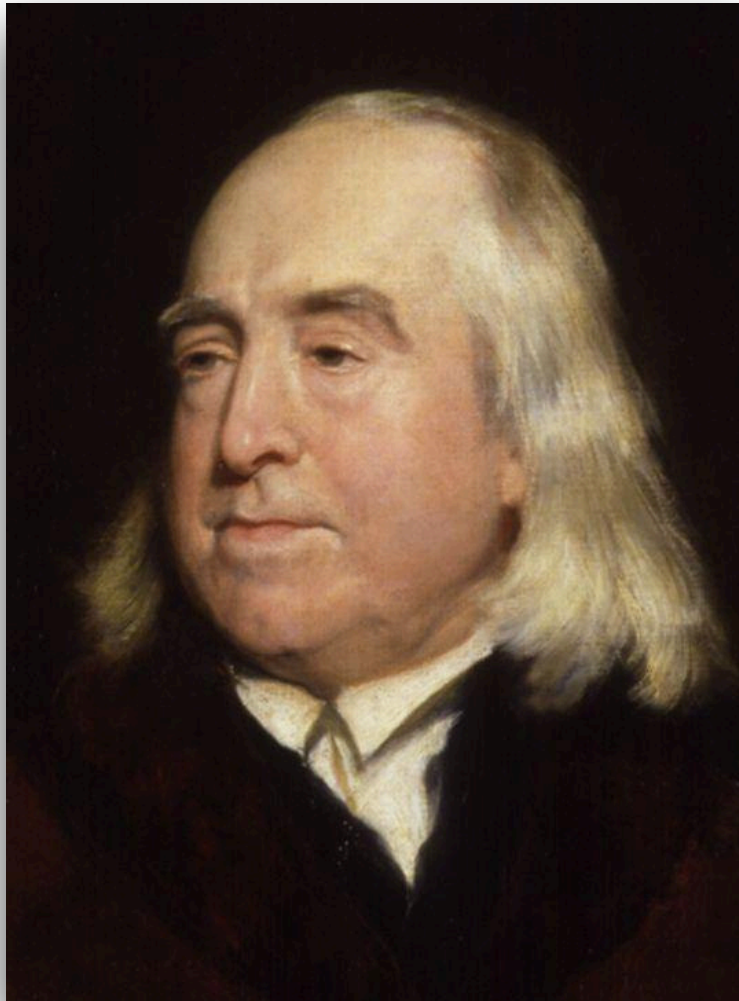
BIG DATA

- WHAT IS BIG DATA?
- HOW IS BIG DATA USED?
- THE GOOD, THE BAD, THE UGLY
- TWO WAYS TO THINK OF BIG DATA
 - PANOPTICON
 - CAPTURE MODEL
- HOW THEN OUGHT WE TO LIVE?

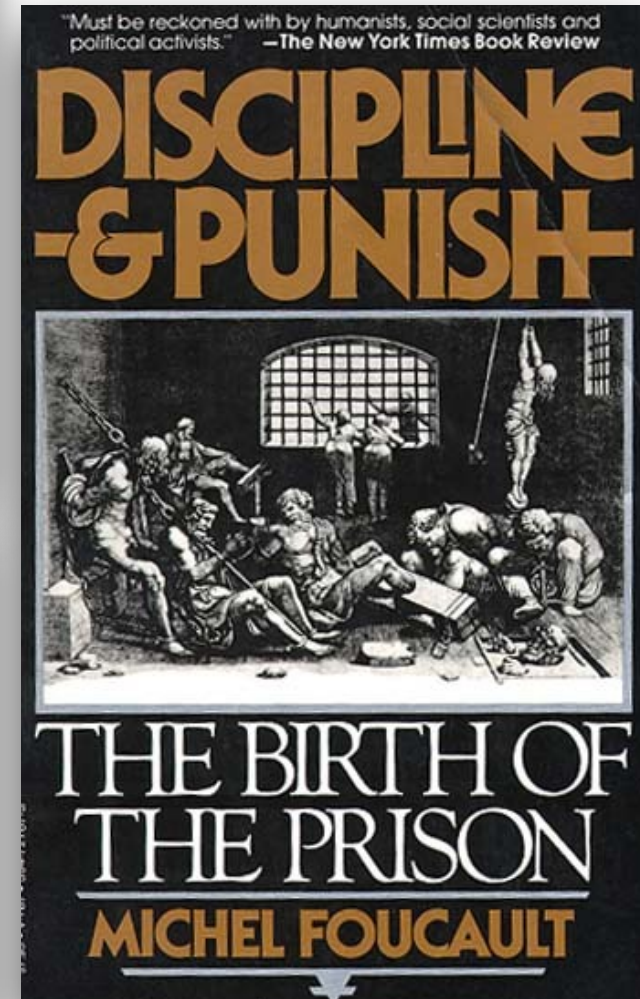
BIG DATA

- WHAT IS BIG DATA?
- HOW IS BIG DATA USED?
- THE GOOD, THE BAD, THE UGLY
- TWO WAYS TO THINK OF BIG DATA
 - PANOPTICON
 - CAPTURE MODEL
- HOW THEN OUGHT WE TO LIVE?

PANOPTICON



Jeremy Bentham

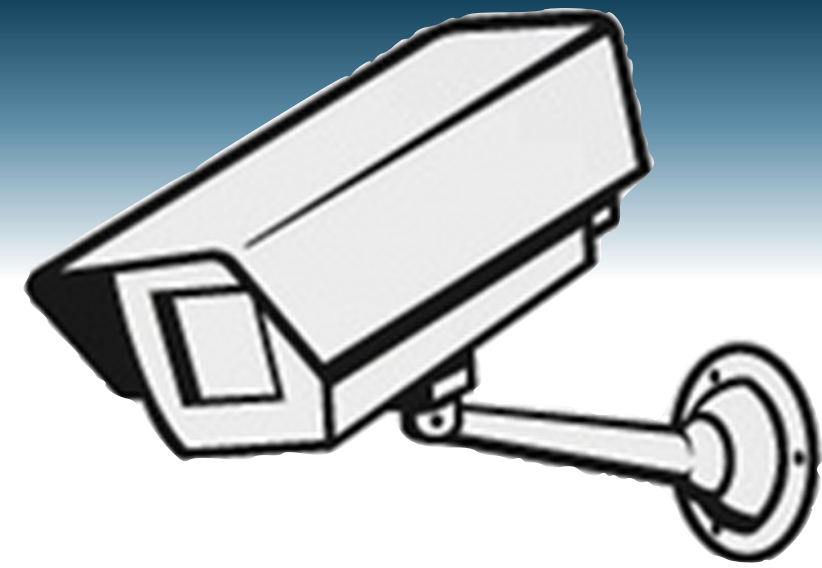


PANOPTICON



PANOPTICON

IN THIS VIEW



- Big Data is about being watched / visual metaphor
- Collecting data is bad for the surveilled
- Collection is non-disruptive
- Data is centralized
- There are “files” on you
- Territorial
 - Not okay to have a camera *here*
- Implies the reorganization of movement to increase or decrease surveillance
 - surveillance is internalized



BIG DATA

- WHAT IS BIG DATA?
- HOW IS BIG DATA USED?
- THE GOOD, THE BAD, THE UGLY
- TWO WAYS TO THINK OF BIG DATA
 - PANOPTICON
 - CAPTURE MODEL
- HOW THEN OUGHT WE TO LIVE?

CAPTURE MODEL

Smart & Final &
Warehouse & Market. Friend & Neighbor.

Welcome To Our Santa Barbara Store
Store #391

See Us On WEB www.smartandfinal.com

Cashier: Leanna Garcia

DATE 03/25/17 TIME 09:35:33

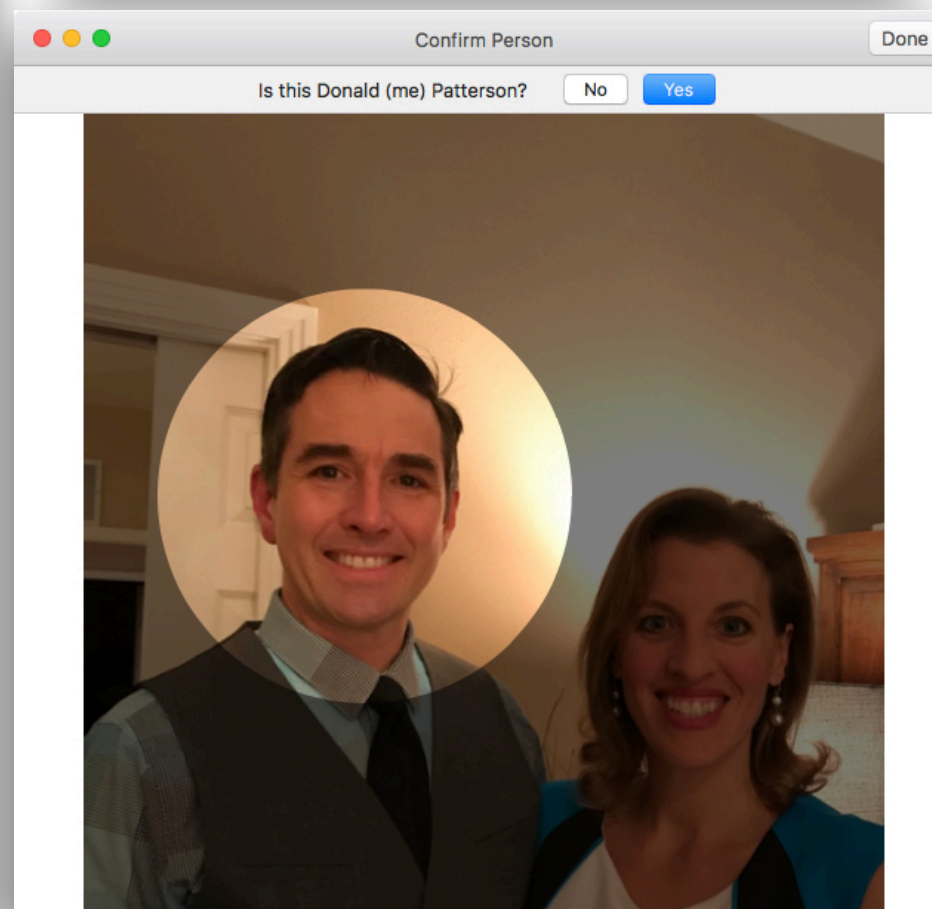
Frito-Lay Big Grab	12.69	F
Barqs Rootbeer	4.00	FD
Was \$5.99 / YOU SAVED ->	\$1.99	
+REDEMPTION	.60	
Fanta Orange	4.00	FD
Was \$5.99 / YOU SAVED ->	\$1.99	
+REDEMPTION	.60	
Fresca Soda	4.00	FD
Was \$5.99 / YOU SAVED ->	\$1.99	
+REDEMPTION	.60	
Mirchi Chkn Ins Lnc	3.89	F
Az Tea Variety Pac	3.99	F
Was \$10.99 / YOU SAVED ->	\$1.00	
+CRV	1.20	
SUBTOTAL	41.57	
Sales Tax	1.07	
TOTAL	42.64	
Visa	TENDER	42.64
Cas Ref# 14	CHANGE	.00
Cash		

TOTAL NUMBER OF ITEMS THIS VISIT--> 3

Smart & Final Store # 391
217 East Gutierrez Street
Santa Barbara, CA 93101
Operator 40008958
DATE 03/25/17 TIME 09:37:14
Account # *****2296
Reference # 187880
APPRVL CODE 025377
Reason Code RMD0
Credit
TOTAL 42.64
BALANCE .00

09:37:14 OP# 40008958 03/25/17
Term:3 Trans # 19 Store # 391

THANK YOU FOR SHOPPING
YOUR SANTA BARBARA SMART AND FINAL
STORE MANAGER: Louis Desales
1 (805) 564-8855



 Associate Professor at Westmont College
Past: University of California, Irvine

 Santa Barbara, CA, United States

 <https://www.djp3.net/>

 Studied Ubiquitous computing at University of Washington
Past: Cornell University and Cornell University

 April 11, 1905

 Lives in Santa Barbara, California
From Washington, District of Columbia

 In a domestic partnership with Sandy Lean Patterson

CAPTURE MODEL

IN THIS VIEW

- Big Data is about computational understanding
- Trying to understand what are meaningful events
- Collecting data is good for the surveilled
- Data must be parsed, synthesized and reconstructed
- Data is distributed
- Correlating data is the goal
 - Encouraged to “Link your profile to Twitter”
- Surveillance is piggy-backed and it grows
- Implies the reorganization of work to increase efficiency and further capture
 - Your legibility to the system is internalized



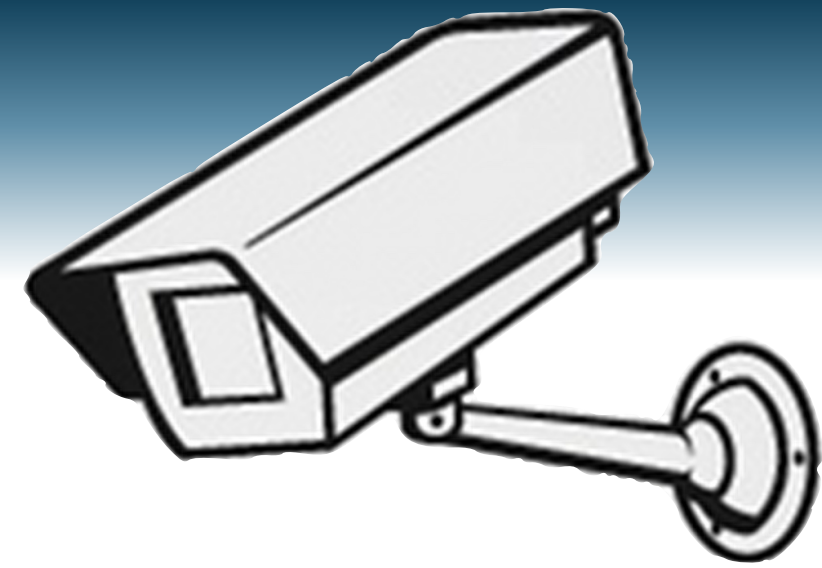




PANOPTICON OR CAPTURE ?

<https://flic.kr/p/qSqFYA>

WHAT ARE THE RISKS?



PANOPTICON

- Ideological/Social Freedom is suppressed
 - Social diversity is our best source of resilience to future crises
- We lose our reasonable ability to control our self-presentation to others
- We lose our ability for political dissent



WHAT ARE THE RISKS?



WHAT ARE THE RISKS?

CAPTURE MODEL

- Intellectual/Imaginative freedom is suppressed
 - Every physical move speaks a grammar of action which triggers a cascade of digital events
 - which cost money
 - communicate to other systems
 - You no longer have any space to think
 - You can't read a book without generating a trail
 - You can't go places without impacting your citizen score etc.
- Every action can be monetized



©2015 Google Culture



BIG DATA

- WHAT IS BIG DATA?
- HOW IS BIG DATA USED?
- THE GOOD, THE BAD, THE UGLY
- TWO WAYS TO THINK OF BIG DATA
 - PANOPTICON
 - CAPTURE MODEL
- HOW THEN OUGHT WE TO LIVE?

THIS WILL NOT BE SOLVED BY TECHNOLOGY

WE MUST ACTIVELY
PURSUE AND SUPPORT
LEGISLATIVE ACTION
TO PUT REASONABLE
LIMITS IN PLACE



WE DON'T WANT TO BE PREY BEING STALKED BY TECHNOLOGY

HOW THEN OUGHT WE TO LIVE?



Submit

**Easy now, but in
the long run you
lose your freedom**

- Disappear
- Rebel
- Lie
- Resist



HOW THEN OUGHT WE TO LIVE?



Submit

Disappear

- Rebel
- Lie
- Resist

**No longer possible.
Between smart phones,
global finance, and
video cameras you can't
reasonably go off-grid
indefinitely**



HOW THEN OUGHT WE TO LIVE?

- 
- Submit
 - Disappear
 - Rebel
 - Lie
 - Resist

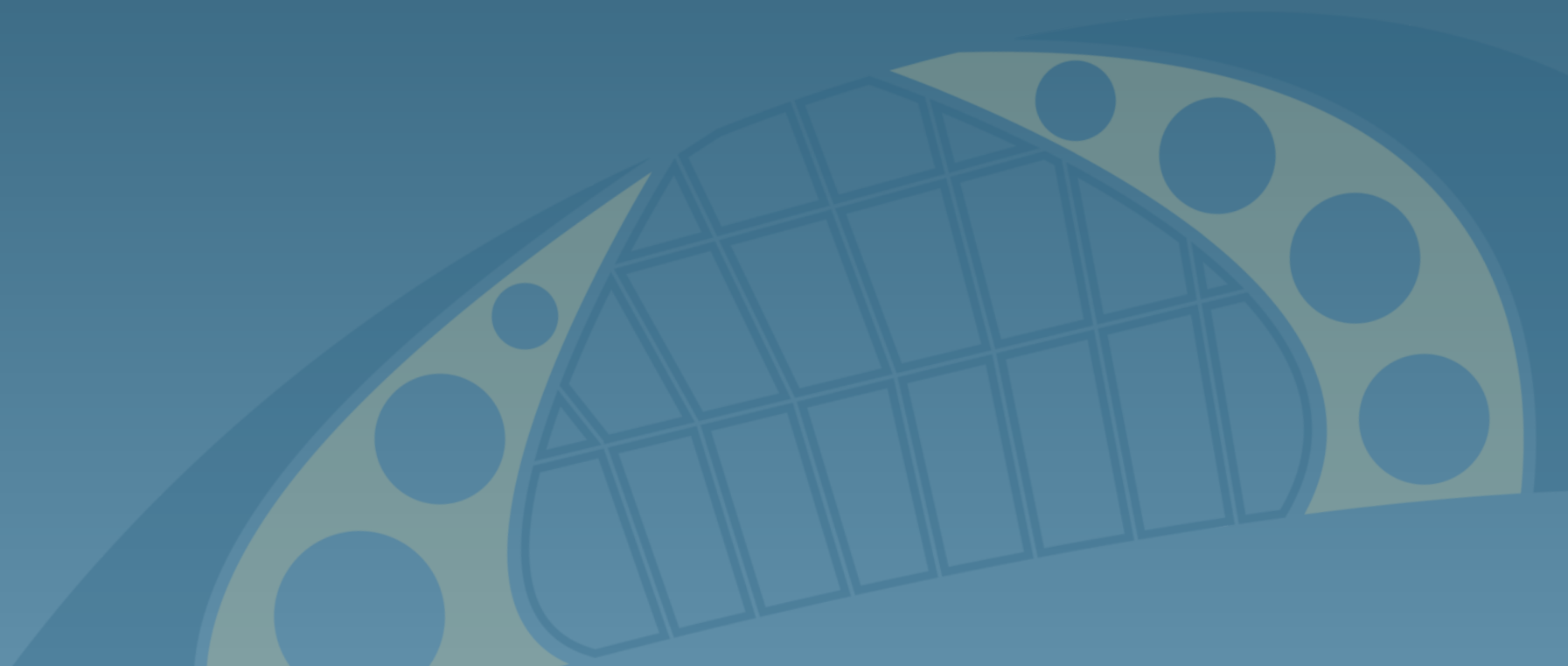
Active law breaking is unlikely to find any sympathy because people rely on these systems to live



LYING CONSIDERED HARMFUL

- It is clear and adamant that the God of the Bible hates lying
 - It is explicitly stated (Proverbs 12:22)
 - It is rooted in corruption and aligns you with evil
 - The fear of losing personal esteem
 - The fear of physical harm (Peter's disowning Jesus)
 - #Jesus
- However, "it is possible to be a person who fears the Lord, walks by faith, and yet feel constrained in extreme, life-threatening situations to oppose evil by lying intentionally." -John Piper
 - Protection of the powerless
 - Rahab, Egyptian midwives
 - Protection of Jews during the Holocaust

IS LYING TO A
DATABASE THE SAME
AS LYING TO A
PERSON?



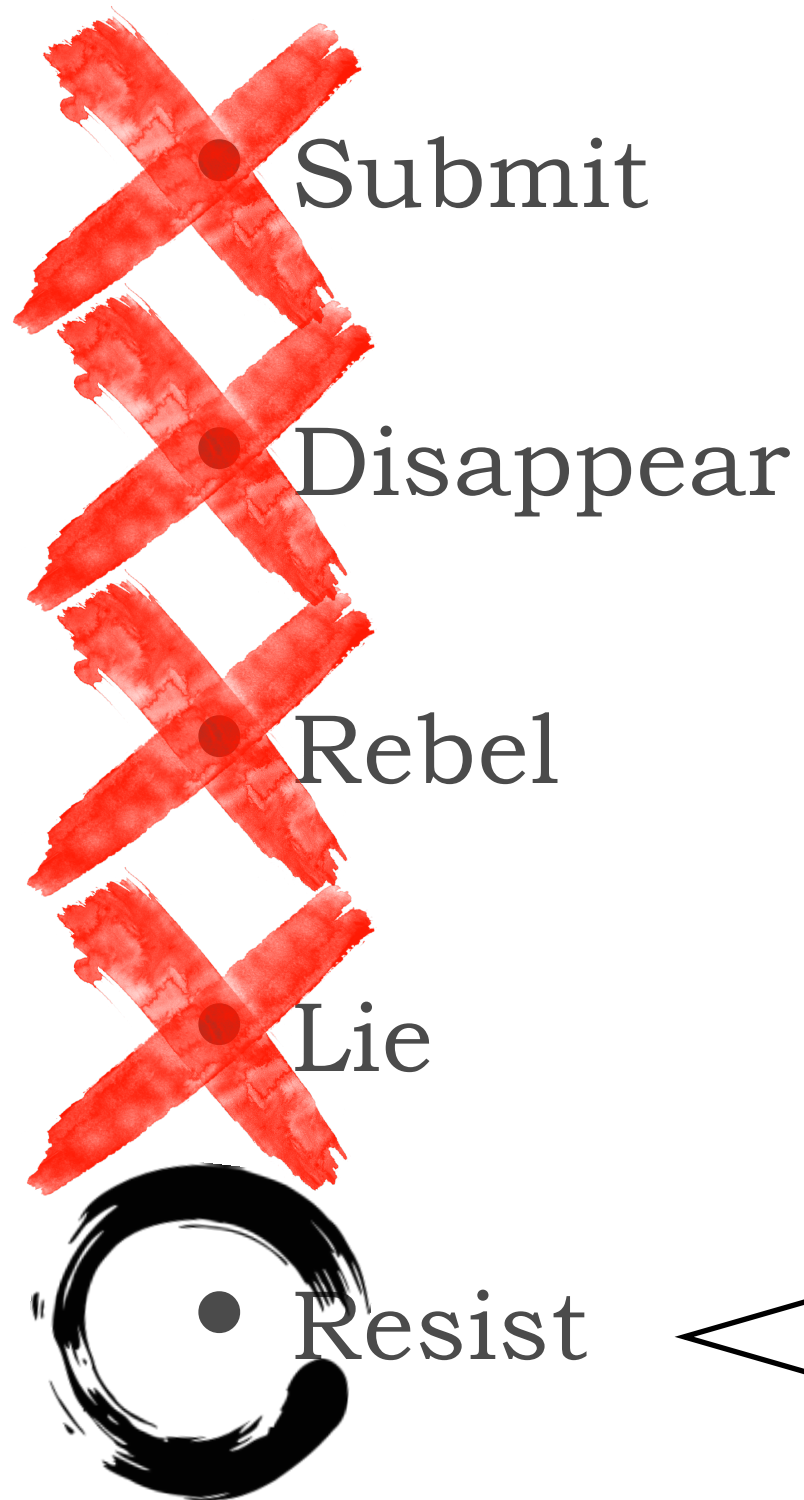
HOW THEN OUGHT WE TO LIVE?

- Submit
- Disappear
- Rebel
- Lie
- Resist

**Extremely difficult to
justify**



HOW THEN OUGHT WE TO LIVE?



**Opt-out when you can
Mess with your data when
you can't
Don't make it easy**

RESISTANCE - DESKTOP COMPUTING

- Install Browser Extensions
 - HTTPS Everywhere
 - resists ISP/Network monitoring of web work
 - Privacy Badger
 - resists website monitoring of you via cookies
 - TrackMeNot
 - resists profiling based on web search behavior
- Turn on “Do not track” on your browser
- Consider a VPN
 - Resists ISP visibility of any network action



RESISTANCE - MOBILE COMPUTING

- Use Signal app for communications
 - Resists monitoring of mobile communications
- Use a VPN
 - Prevents your cell-phone company from monitoring your network behavior
- Turn off all application access to
 - location
 - microphone
 - contacts
 - camera
 - except those that you know critically need it



RESISTANCE - EMBODIED COMPUTING

- Send your phone on a holiday
 - Turn it off
 - Let other people carry it when you don't need it
 - Create plausible deniability
- Physically block your cameras
- Mess with loyalty cards
 - Use someone else's phone number at the store
 - Trade Plenti cards
 - Register a Starbucks card on your account and then trade it with someone else
- Use incognito mode on your browser (protects against people who get access to your computer)



RESISTANCE - HAVE FUN WITH IT

- Record and play fake conversations when you aren't around and see if advertising changes
 - Resists OK Google, Siri, Amazon Echo
- Dirty your data
 - Go on a fake vacation on Facebook
 - Change your demographics on Social Media
- Celebrate the 4th of July
 - by deleting all of your web browser cookies for the year all at once
- Misspell your name on forms and track how your information is traded



FOR FURTHER READING

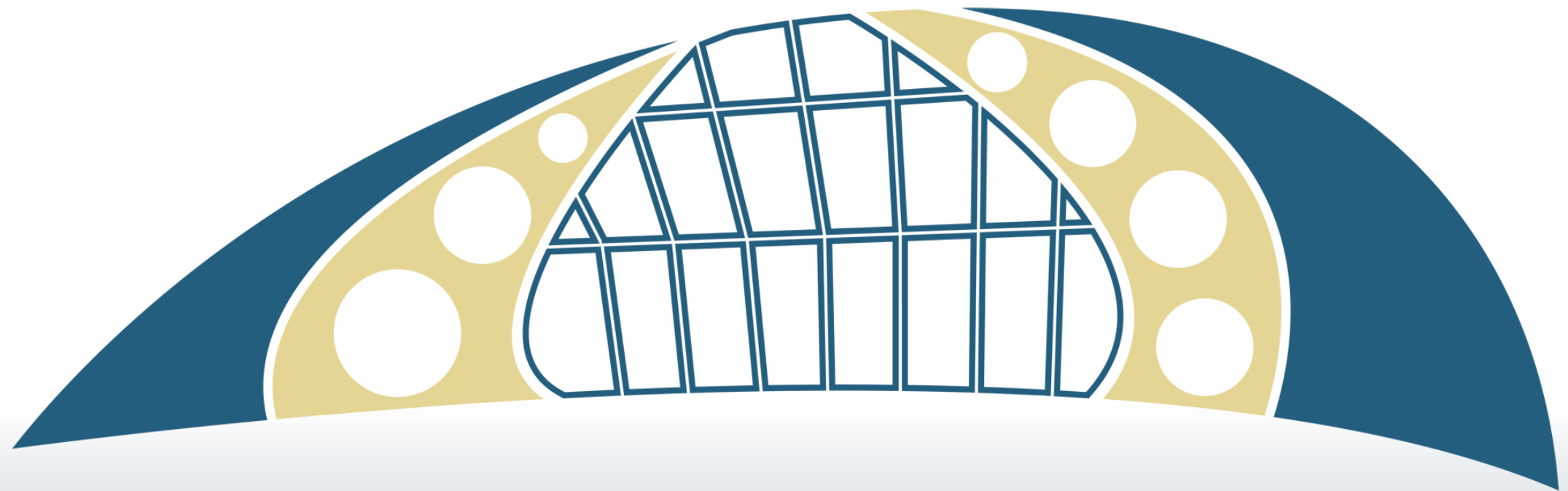


- [illegible]

FOR FURTHER READING

- EPIC: Privacy and Consumer Profiling
 - <https://epic.org/privacy/profiling/>





WESTMONT **INSPIRED**
— COMPUTING LAB —

THANK YOU!

HOW ARE YOU TRACKED

- Cell phone super cookies
- Toll booth transponder
- Cell phone connections to towers
- Wifi connection
- Shopper cards
- Credit report
- Facebook Messenger conversations
 - Your facebook emotes
- Your use of web browser
 - Across multiple devices as you log in
- When you swipe a credit card
- When you read a QR code
- License plate



NOTE ON TRACKING PEOPLE VS OBJECTS

- Objects are often proxies for people
- We have models of representations that fail short of reality
- context is lost and data that is captured may present a picture of the world which is not as the world actually is
 - This is a risk



THE SELFIE AS A RESPONSE TO SURVEILLANCE

- A Selfie is self-surveillance
 - high volume
 - controlled
 - presents a specific view of self
 - that swamps the automatically curated and publicized data of self



THE HUMAN COMPUTER INTERACTION IS LOST

- The capture model erodes the idea that we are interacting with a computer at defined moments.
 - Moments which we used to choose
 - Moments which we used to be able to opt-out of
 - Moments for which we understood the implications



WHAT'S THE POINT

- Privacy isn't just about keeping our data to ourselves
 - It is also about our existence as social beings and our desire, perhaps right, to control our presentation and decide what our identity is



HIDING

- Anonymous really isn't

<https://arstechnica.com/tech-policy/2009/09/your-secrets-live-online-in-databases-of-ruin/>

