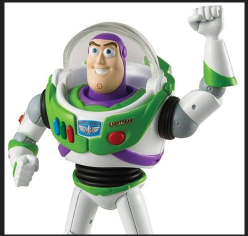
data: past, present, and

&

jupyter

&



chris wiggins + matt jones

Lec 1, 2023 ; data-ppf.gith

Let’s start with a data story…

about how you are valued

(by the machines)

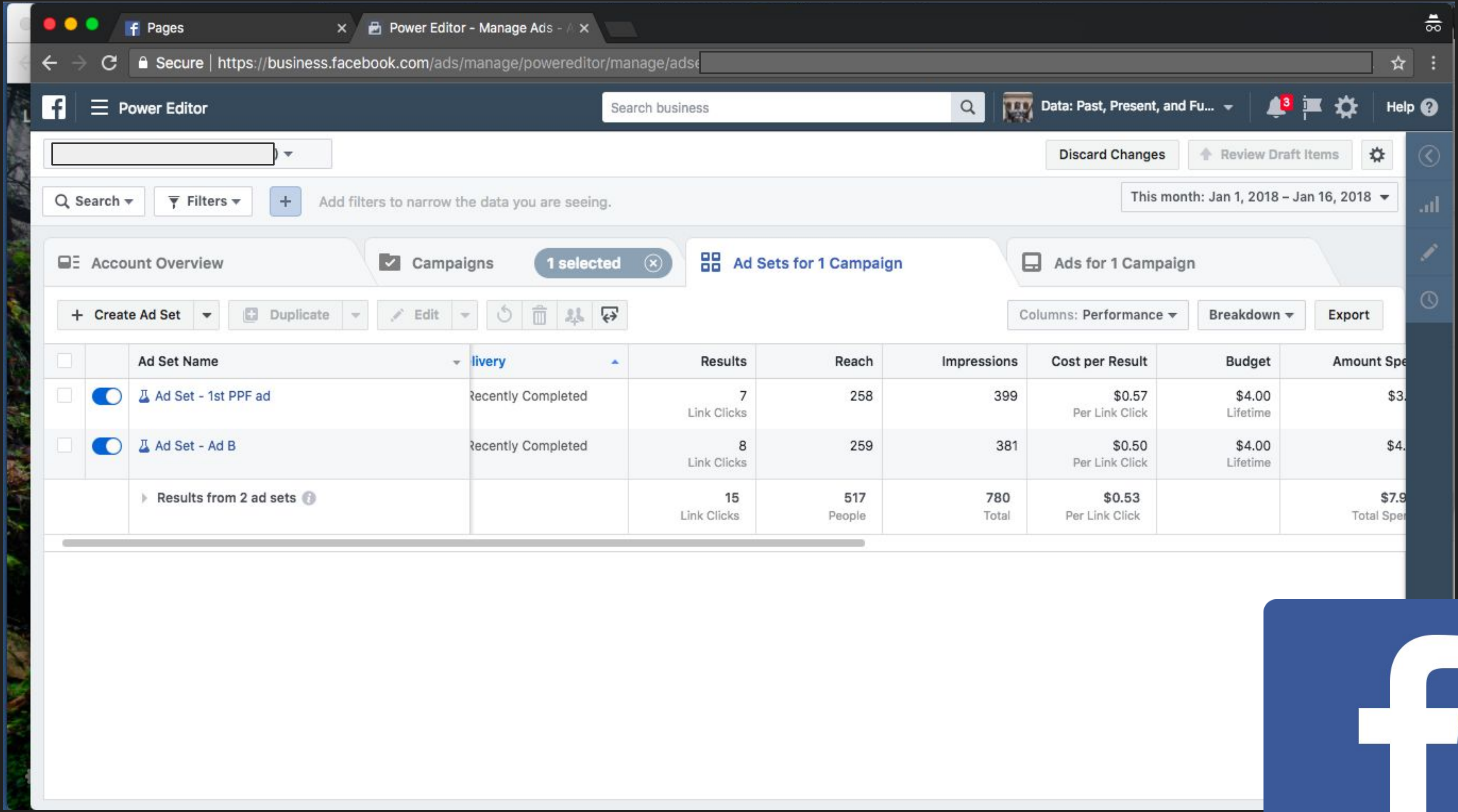
Suggested Page

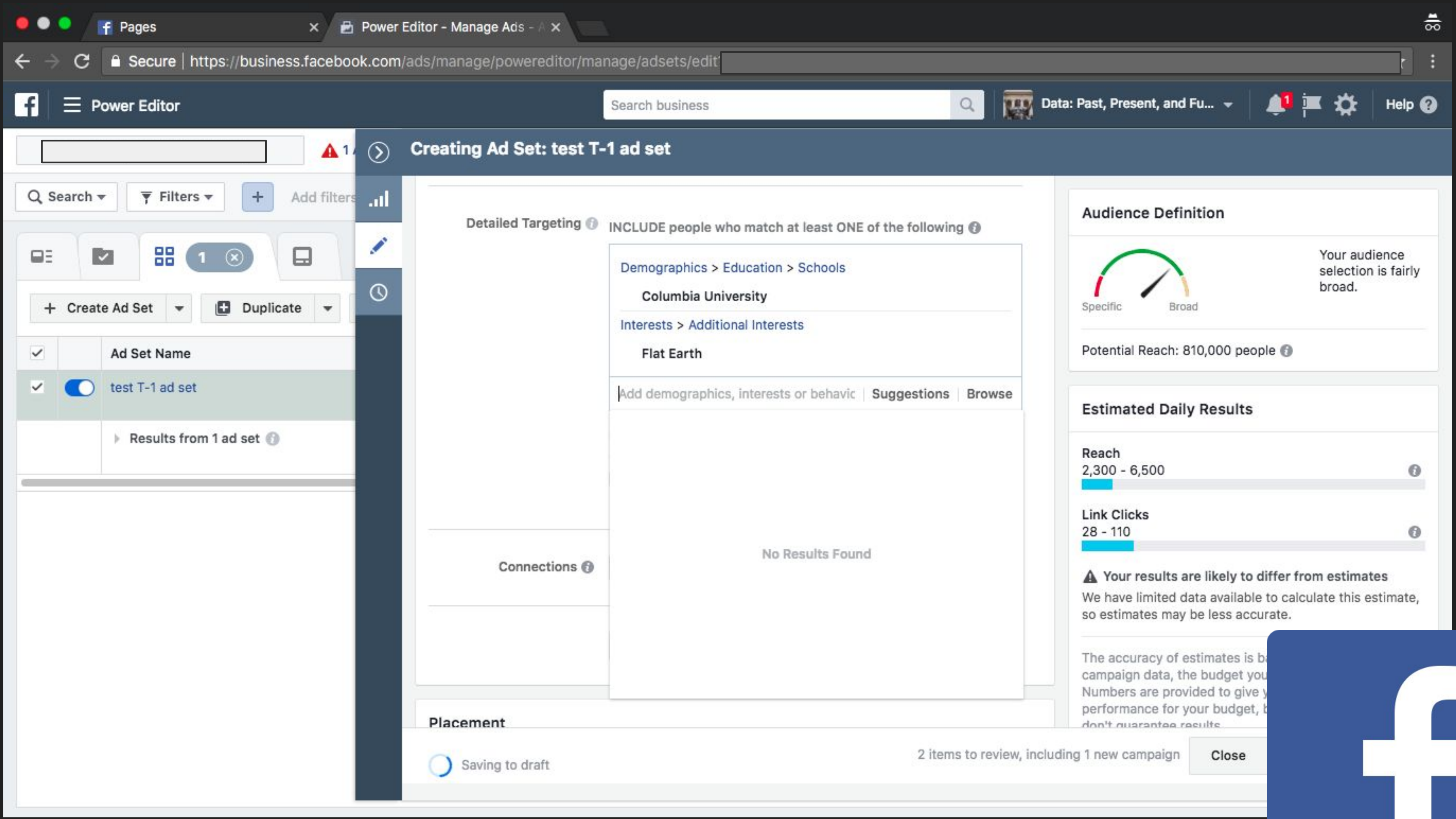
Data: Past, Present, and Future
Sponsored

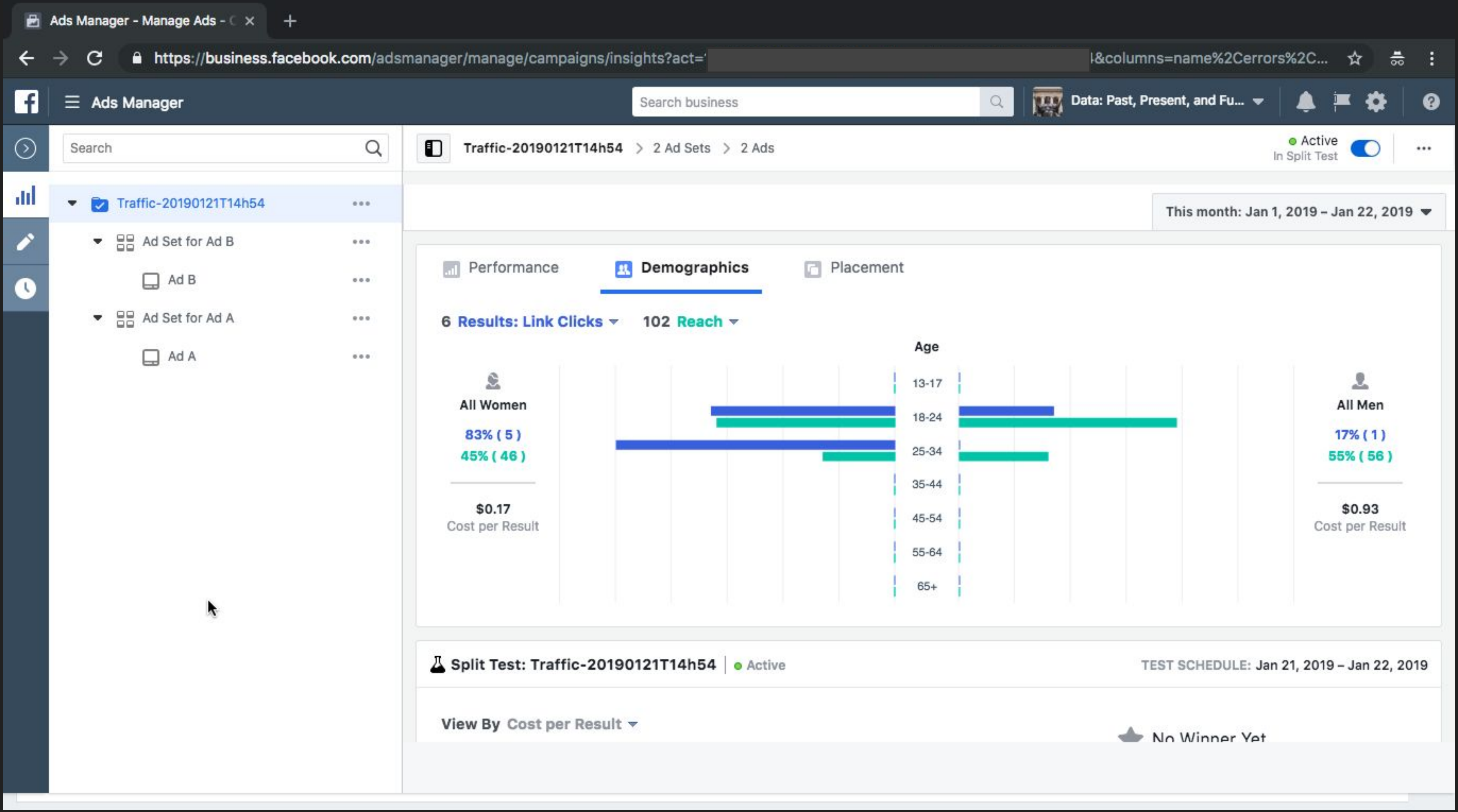
how did this end up in my news feed? find out in new course:
https://data-ppf.github.io/

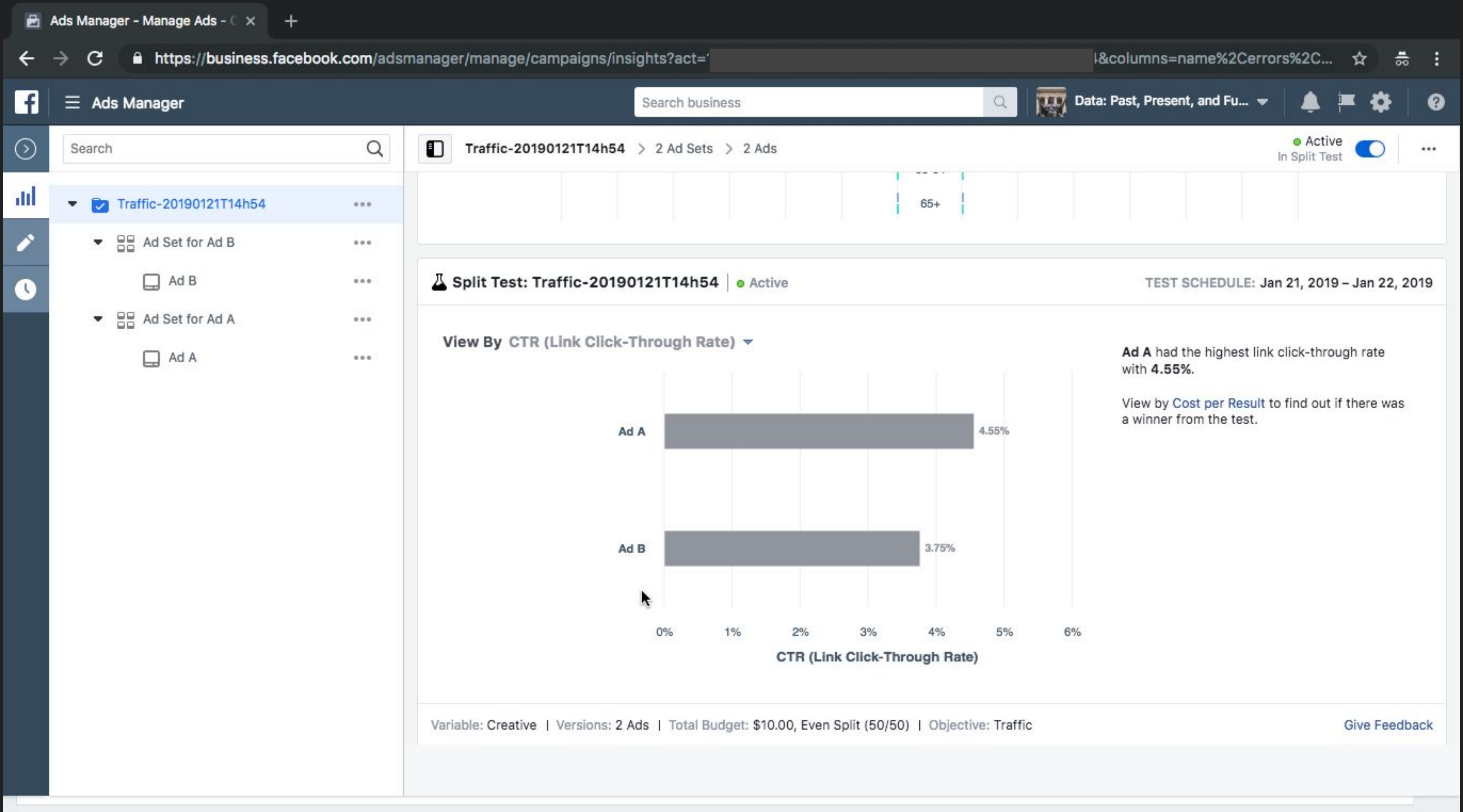
Data: Past, Present, and Future
Community
3 people like this

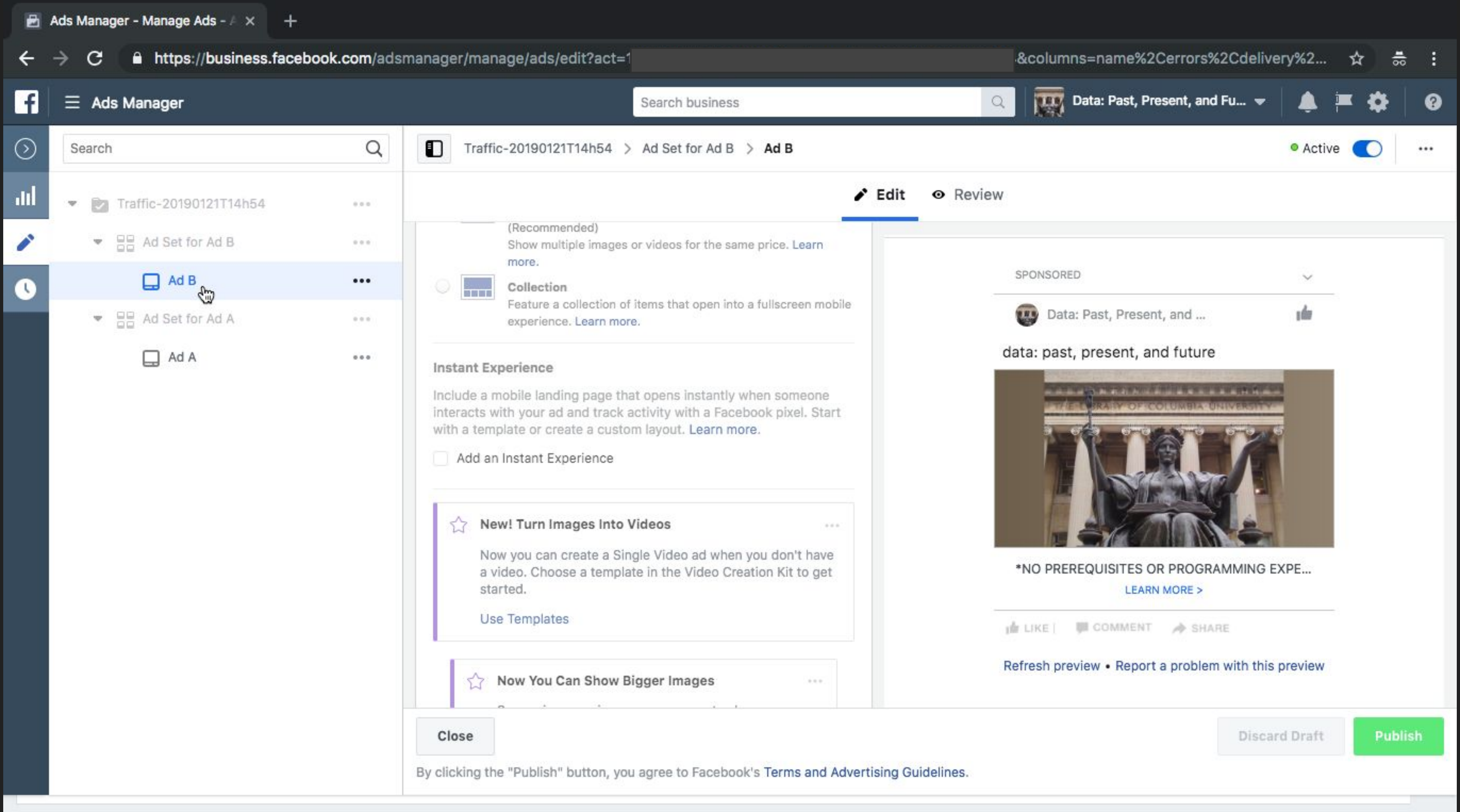
Like Page

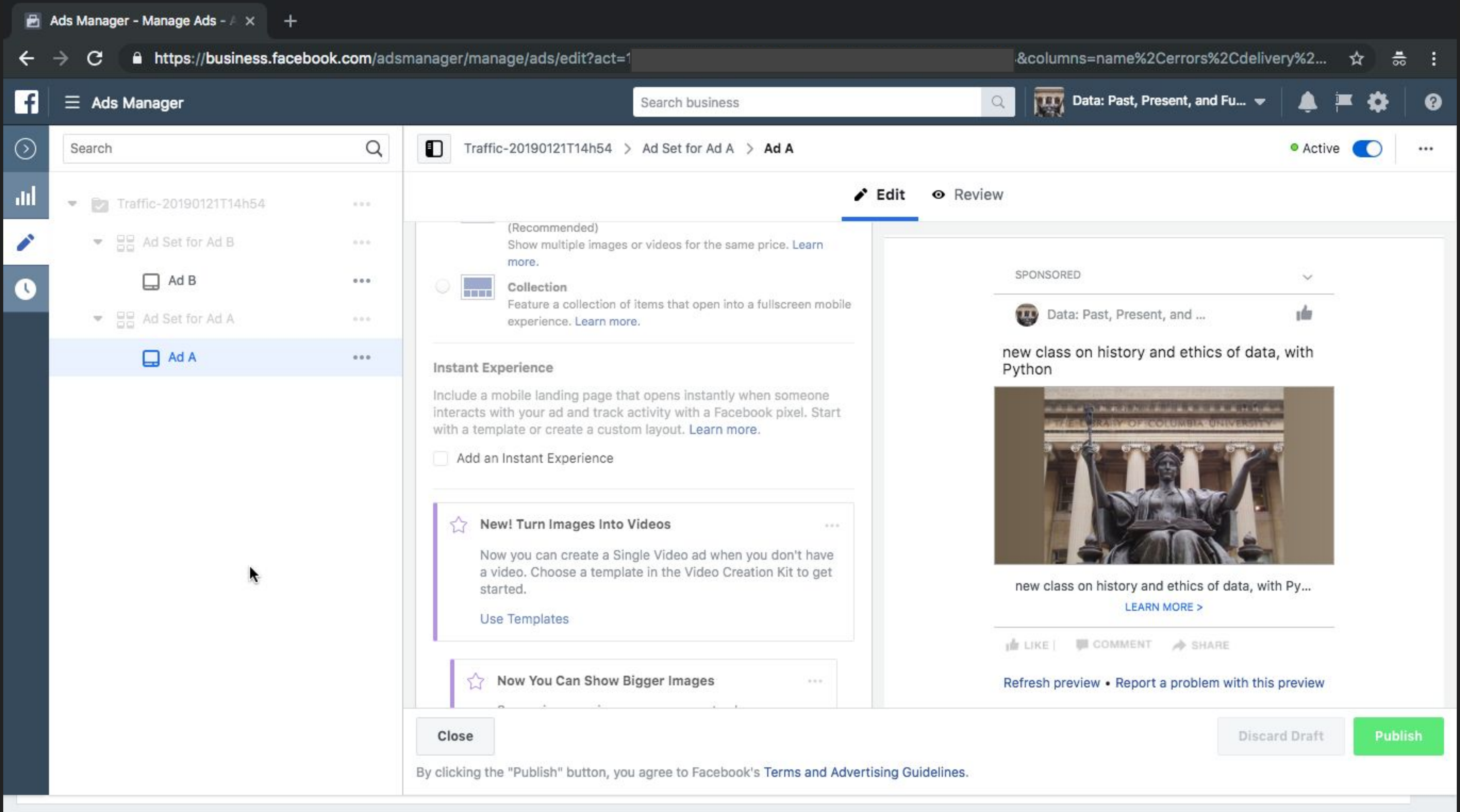

Pages
Power Editor - Manage Ads
Secure
https://business.facebook.com/ads/manage/powereditor/manage/adse [redacted]
Power Editor
Search business
Data: Past, Present, and Fu...
Help
Discard Changes
Review Draft Items
Search
Filters
Add filters to narrow the data you are seeing.
This month: Jan 1, 2018 – Jan 16, 2018
Account Overview
Campaigns
1 selected
Ad Sets for 1 Campaign
Ads for 1 Campaign
Create Ad Set
Duplicate
Edit
Columns: Performance
Breakdown
Export
Ad Set Name
livery
Results
Reach
Impressions
Cost per Result
Budget
Amount Spe
Ad Set - 1st PPF ad
Recently Completed
7 Link Clicks
258
399
$0.57 Per Link Click
$4.00 Lifetime
$3.
Ad Set - Ad B
Recently Completed
8 Link Clicks
259
381
$0.50 Per Link Click
$4.00 Lifetime
$4.
Results from 2 ad sets
15 Link Clicks
517 People
780 Total
$0.53 Per Link Click
$7.9 Total Spe


Pages
Power Editor - Manage Ads
Secure | https://business.facebook.com/ads/manage/powereditor/manage/adsets/edit
Power Editor
Search business
Data: Past, Present, and Fu...
Help
Search
Filters
Add filters
Create Ad Set
Duplicate
Ad Set Name
test T-1 ad set
Results from 1 ad set
Creating Ad Set: test T-1 ad set
Detailed Targeting
INCLUDE people who match at least ONE of the following
Demographics > Education > Schools
Columbia University
Interests > Additional Interests
Flat Earth
Add demographics, interests or behavic
Suggestions
Browse
No Results Found
Connections
Placement
Saving to draft
2 items to review, including 1 new campaign
Close
Audience Definition
Your audience selection is fairly broad.
Specific
Broad
Potential Reach: 810,000 people
Estimated Daily Results
Reach
2,300 - 6,500
Link Clicks
28 - 110
Your results are likely to differ from estimates
We have limited data available to calculate this estimate, so estimates may be less accurate.

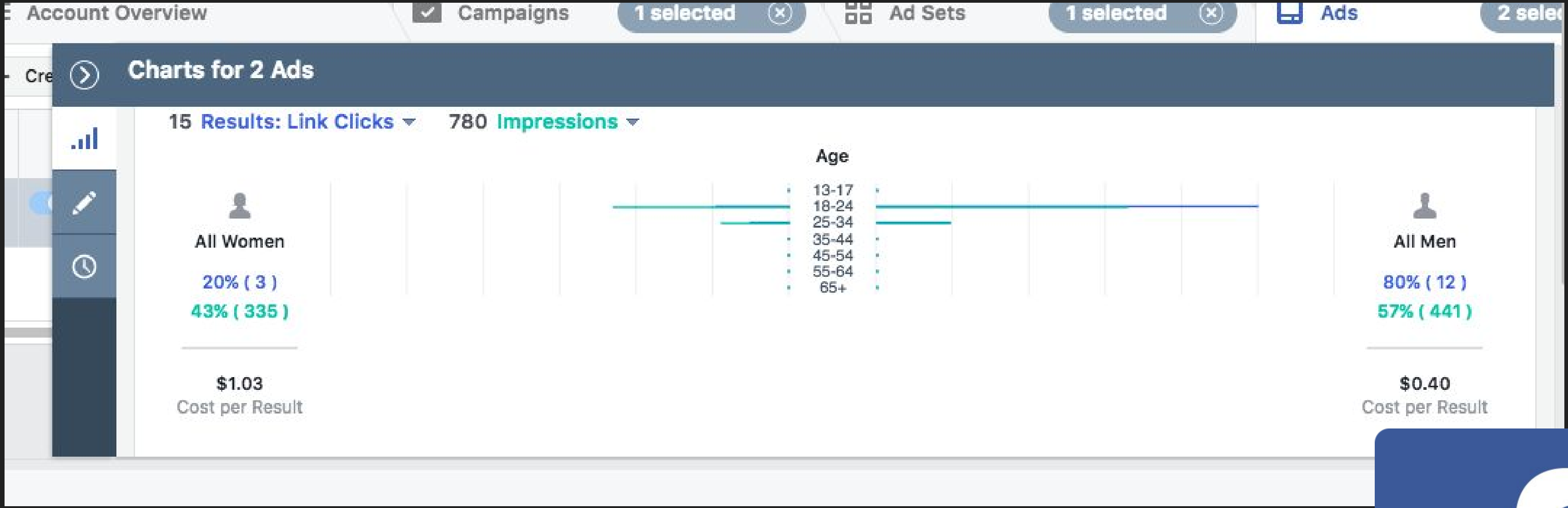

Ads Manager - Manage Ads - (
https://business.facebook.com/adsmanager/manage/campaigns/insights?act=' [redacted] &columns=name%2Cerrors%2C...
Ads Manager
Search business
Data: Past, Present, and Fu...
Search
Traffic-20190121T14h54
Ad Set for Ad B
Ad B
Ad Set for Ad A
Ad A
Traffic-20190121T14h54 > 2 Ad Sets > 2 Ads
Active
In Split Test
This month: Jan 1, 2019 – Jan 22, 2019
Performance
Demographics
Placement
6 Results: Link Clicks
102 Reach
Age
All Women
83% (5)
45% (46)
$0.17
Cost per Result
13-17
18-24
25-34
35-44
45-54
55-64
65+
All Men
17% (1)
55% (56)
$0.93
Cost per Result
Split Test: Traffic-20190121T14h54
Active
TEST SCHEDULE: Jan 21, 2019 – Jan 22, 2019
View By Cost per Result
No Winner Yet


Ads Manager - Manage Ads - (
https://business.facebook.com/adsmanager/manage/campaigns/insights?act='[redacted]&columns=name%2Cerrors%2C...
Ads Manager
Search business
Data: Past, Present, and Fu...
Search
Traffic-20190121T14h54
Ad Set for Ad B
Ad B
Ad Set for Ad A
Ad A
Traffic-20190121T14h54 > 2 Ad Sets > 2 Ads
Active
In Split Test
65+
Split Test: Traffic-20190121T14h54 | Active
TEST SCHEDULE: Jan 21, 2019 – Jan 22, 2019
View By CTR (Link Click-Through Rate)
Ad A had the highest link click-through rate with 4.55%.
View by Cost per Result to find out if there was a winner from the test.
Ad A
4.55%
Ad B
3.75%
0%
1%
2%
3%
4%
5%
6%
CTR (Link Click-Through Rate)
Variable: Creative | Versions: 2 Ads | Total Budget: $10.00, Even Split (50/50) | Objective: Traffic
Give Feedback


Ads Manager - Manage Ads - A
https://business.facebook.com/adsmanager/manage/ads/edit?act=1[redacted]&columns=name%2Cerrors%2Cdelivery%2...
Ads Manager
Search business
Data: Past, Present, and Fu...
Search
Traffic-20190121T14h54
Ad Set for Ad B
Ad B
Ad Set for Ad A
Ad A
Traffic-20190121T14h54 > Ad Set for Ad B > Ad B
Active
Edit
Review
(Recommended)
Show multiple images or videos for the same price. Learn more.
Collection
Feature a collection of items that open into a fullscreen mobile experience. Learn more.
Instant Experience
Include a mobile landing page that opens instantly when someone interacts with your ad and track activity with a Facebook pixel. Start with a template or create a custom layout. Learn more.
Add an Instant Experience
New! Turn Images Into Videos
Now you can create a Single Video ad when you don't have a video. Choose a template in the Video Creation Kit to get started.
Use Templates
Now You Can Show Bigger Images
SPONSORED
Data: Past, Present, and ...
data: past, present, and future
*NO PREREQUISITES OR PROGRAMMING EXPE...
LEARN MORE >
LIKE
COMMENT
SHARE
Refresh preview • Report a problem with this preview
Close
Discard Draft
Publish
By clicking the "Publish" button, you agree to Facebook's Terms and Advertising Guidelines.


Ads Manager - Manage Ads - A
https://business.facebook.com/adsmanager/manage/ads/edit?act=1[redacted]&columns=name%2Cerrors%2Cdelivery%2...
Ads Manager
Search business
Data: Past, Present, and Fu...
Search
Traffic-20190121T14h54
Ad Set for Ad B
Ad B
Ad Set for Ad A
Ad A
Traffic-20190121T14h54 > Ad Set for Ad A > Ad A
Active
Edit
Review
(Recommended)
Show multiple images or videos for the same price. Learn more.
Collection
Feature a collection of items that open into a fullscreen mobile experience. Learn more.
Instant Experience
Include a mobile landing page that opens instantly when someone interacts with your ad and track activity with a Facebook pixel. Start with a template or create a custom layout. Learn more.
Add an Instant Experience
New! Turn Images Into Videos
Now you can create a Single Video ad when you don't have a video. Choose a template in the Video Creation Kit to get started.
Use Templates
Now You Can Show Bigger Images
SPONSORED
Data: Past, Present, and ...
new class on history and ethics of data, with Python
new class on history and ethics of data, with Py...
LEARN MORE >
LIKE
COMMENT
SHARE
Refresh preview • Report a problem with this preview
Close
Discard Draft
Publish
By clicking the "Publish" button, you agree to Facebook's Terms and Advertising Guidelines.

2 0 1 8



Account Overview
Campaigns
1 selected
Ad Sets
1 selected
Ads
Charts for 2 Ads
15 Results: Link Clicks
780 Impressions
Age
13-17
18-24
25-34
35-44
45-54
55-64
65+
All Women
20% (3)
43% (335)
$1.03
Cost per Result
All Men
80% (12)
57% (441)
$0.40
Cost per Result

Location	Impressions	Spend	Results
Total for account	**12,940**	**$25.00**	–
Brazil	38	$0.06	–
Indonesia	5,638	$3.38	–
Bosnia and Herzegovina	3	$0.01	–
Ukraine	2	$0.01	–
Trinidad and Tobago	1	$0.00	–
Greece	2	$0.00	–
Kuwait	18	$0.00	–

Objective: All × Status: All ×

Spend
$25.00

Jan 11

Campaigns Ad Groups

Name
Summary for 1 item
On Twitter
On Twitter Audience Platform
data:PPF spring 2018 Website clicks or conversions
On Twitter
On Twitter Audience Platform

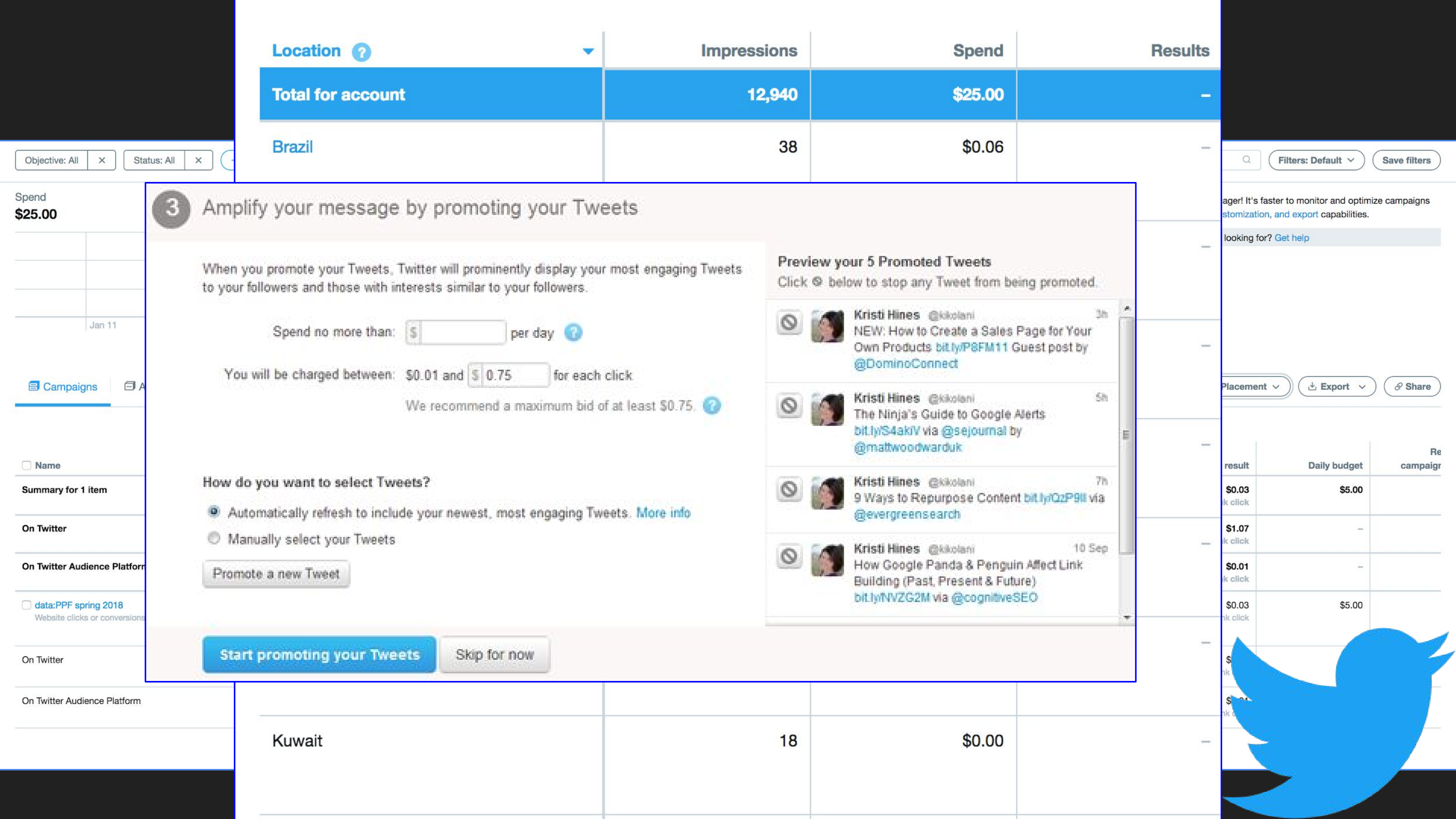
Filters: Default Save filters

ager! It's faster to monitor and optimize campaigns
istomization, and export capabilities.

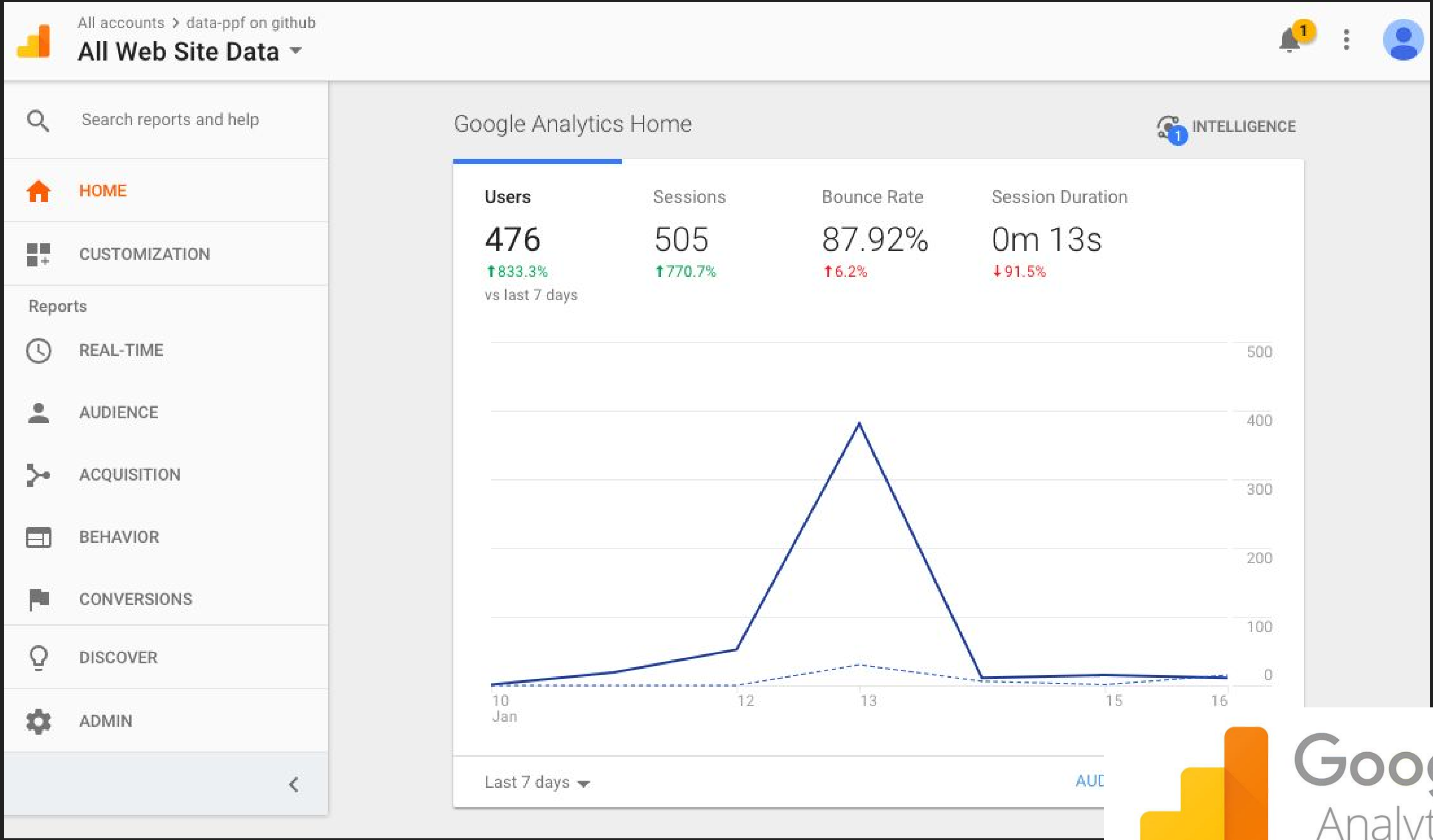
looking for? Get help

Placement Export Share

result	Daily budget	Re campaign
$0.03 nk click	$5.00	
$1.07 nk click	–	
$0.01 nk click	–	
$0.03 nk click	$5.00	



Location
Impressions
Spend
Results
Total for account
12,940
$25.00
Brazil
38
$0.06
Kuwait
18
$0.00
Objective: All
Status: All
Spend
$25.00
Jan 11
Campaigns
Name
Summary for 1 item
On Twitter
On Twitter Audience Platform
data:PPF spring 2018
Website clicks or conversions
On Twitter
On Twitter Audience Platform
Filters: Default
Save filters
ager! It's faster to monitor and optimize campaigns
stomization, and export capabilities.
looking for? Get help
Placement
Export
Share
result
Daily budget
$0.03
$5.00
$1.07
$0.01
$0.03
$5.00
3
Amplify your message by promoting your Tweets
When you promote your Tweets, Twitter will prominently display your most engaging Tweets to your followers and those with interests similar to your followers.
Spend no more than: $ per day
You will be charged between: $0.01 and $ 0.75 for each click
We recommend a maximum bid of at least $0.75.
How do you want to select Tweets?
Automatically refresh to include your newest, most engaging Tweets. More info
Manually select your Tweets
Promote a new Tweet
Start promoting your Tweets
Skip for now
Preview your 5 Promoted Tweets
Click below to stop any Tweet from being promoted.
Kristi Hines @kikolani 3h
NEW: How to Create a Sales Page for Your Own Products bit.ly/P8FM11 Guest post by @DominoConnect
Kristi Hines @kikolani 5h
The Ninja's Guide to Google Alerts bit.ly/S4akiV via @sejournal by @mattwoodwarduk
Kristi Hines @kikolani 7h
9 Ways to Repurpose Content bit.ly/QzP9II via @evergreensearch
Kristi Hines @kikolani 10 Sep
How Google Panda & Penguin Affect Link Building (Past, Present & Future) bit.ly/NVZG2M via @cognitiveSEO

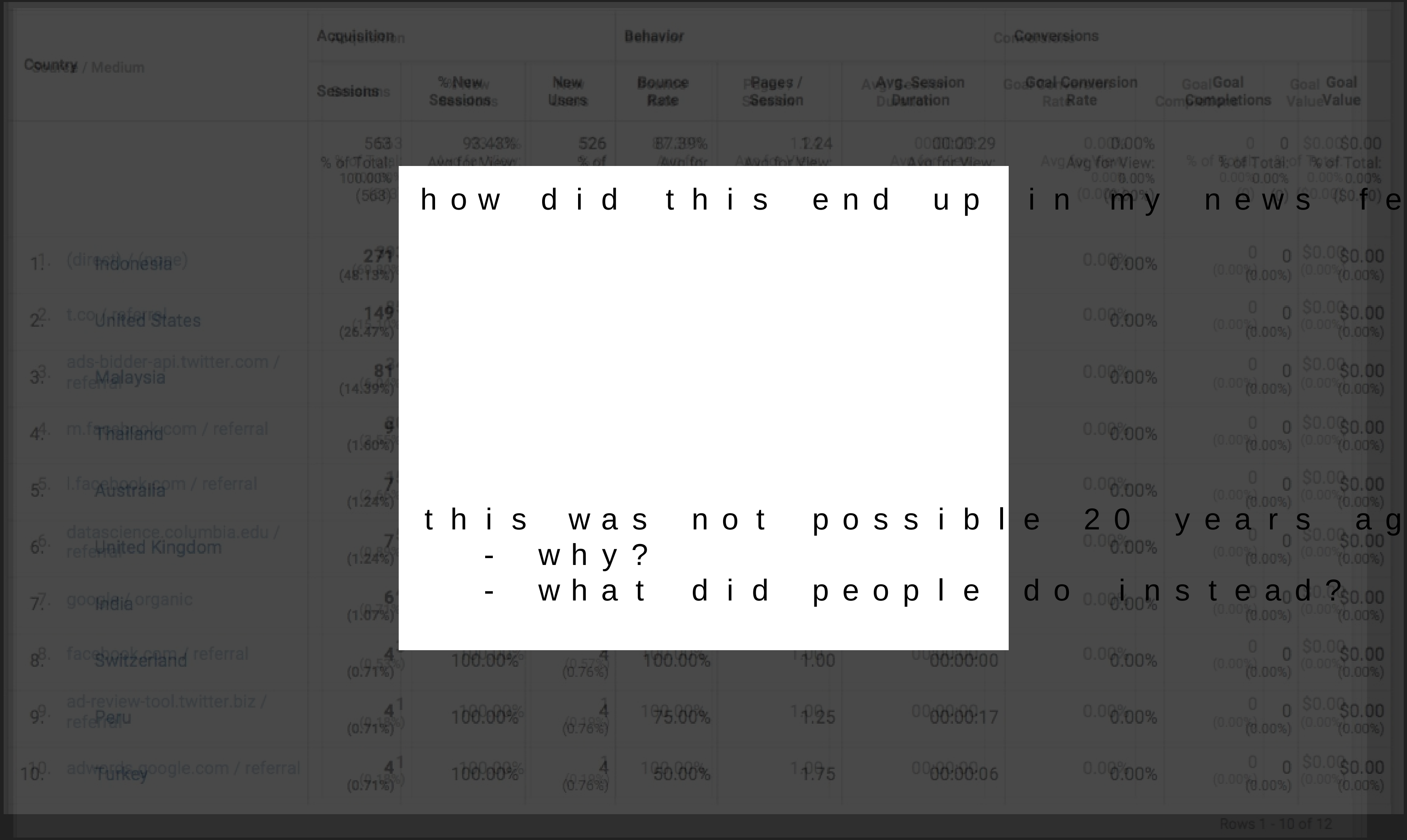


Google Analytics

Country		Acquisition			Behavior			Conversions		
		Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration	Goal Conversion Rate	Goal Completions	Goal Value
		563 % of Total: 100.00% (563)	93.43% Avg for View: 93.43% (0.00%)	526 % of Total: 100.00% (526)	87.39% Avg for View: 87.39% (0.00%)	1.24 Avg for View: 1.24 (0.00%)	00:00:29 Avg for View: 00:00:29 (0.00%)	0.00% Avg for View: 0.00% (0.00%)	0 % of Total: 0.00% (0)	$0.00 % of Total: 0.00% ($0.00)
1.	Indonesia	**271** (48.13%)	95.20%	258 (49.05%)	92.99%	1.17	00:00:06	0.00%	0 (0.00%)	$0.00 (0.00%)
2.	United States	**149** (26.47%)	85.23%	127 (24.14%)	80.54%	1.40	00:01:33	0.00%	0 (0.00%)	$0.00 (0.00%)
3.	Malaysia	**81** (14.39%)	98.77%	80 (15.21%)	87.65%	1.17	<00:00:01	0.00%	0 (0.00%)	$0.00 (0.00%)
4.	Thailand	**9** (1.60%)	100.00%	9 (1.71%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	$0.00 (0.00%)
5.	Australia	**7** (1.24%)	85.71%	6 (1.14%)	42.86%	2.00	00:00:04	0.00%	0 (0.00%)	$0.00 (0.00%)
6.	United Kingdom	**7** (1.24%)	100.00%	7 (1.33%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	$0.00 (0.00%)
7.	India	**6** (1.07%)	100.00%	6 (1.14%)	50.00%	1.50	00:00:53	0.00%	0 (0.00%)	$0.00 (0.00%)
8.	Switzerland	**4** (0.71%)	100.00%	4 (0.76%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	$0.00 (0.00%)
9.	Peru	**4** (0.71%)	100.00%	4 (0.76%)	75.00%	1.25	00:00:17			
10.	Turkey	**4** (0.71%)	100.00%	4 (0.76%)	50.00%	1.75	00:00:06			

Google Analytics

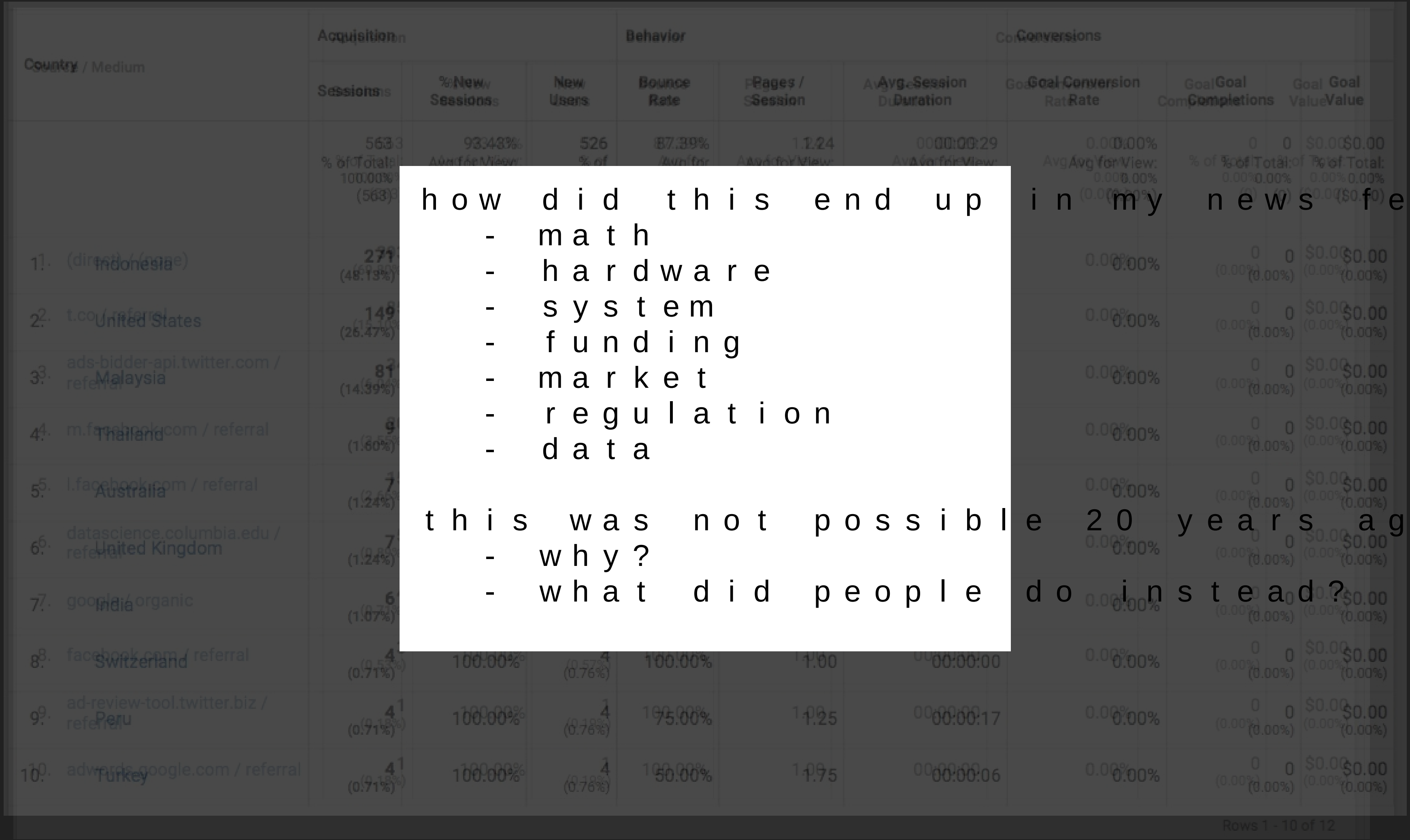
Source / Medium	Acquisition			Behavior			Conversions		
	Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration	Goal Conversion Rate	Goal Completions	Goal Value
	563 % of Total: 100.00% (563)	93.43% Avg for View: 93.43% (0.00%)	526 % of Total: 100.00% (526)	87.39% Avg for View: 87.39% (0.00%)	1.24 Avg for View: 1.24 (0.00%)	00:00:29 Avg for View: 00:00:29 (0.00%)	0.00% Avg for View: 0.00% (0.00%)	0 % of Total: 0.00% (0)	$0.00 % of Total: 0.00% ($0.00)
1. (direct) / (none)	**393** **(69.80%)**	95.93%	377 (71.67%)	89.31%	1.21	00:00:17	0.00%	0 (0.00%)	$0.00 (0.00%)
2. t.co / referral	**85** **(15.10%)**	91.76%	78 (14.83%)	77.65%	1.36	00:00:33	0.00%	0 (0.00%)	$0.00 (0.00%)
3. ads-bidder-api.twitter.com / referral	**34** **(6.04%)**	82.35%	28 (5.32%)	88.24%	1.18	00:00:28	0.00%	0 (0.00%)	$0.00 (0.00%)
4. m.facebook.com / referral	**20** **(3.55%)**	100.00%	20 (3.80%)	95.00%	1.05	00:00:08	0.00%	0 (0.00%)	$0.00 (0.00%)
5. l.facebook.com / referral	**15** **(2.66%)**	93.33%	14 (2.66%)	80.00%	1.80	00:04:34	0.00%	0 (0.00%)	$0.00 (0.00%)
6. datascience.columbia.edu / referral	**5** **(0.89%)**	20.00%	1 (0.19%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	$0.00 (0.00%)
7. google / organic	**4** **(0.71%)**	25.00%	1 (0.19%)	50.00%	1.75	00:06:15	0.00%	0 (0.00%)	$0.00 (0.00%)
8. facebook.com / referral	**3** **(0.53%)**	100.00%	3 (0.57%)	100.00%	1.00	00:00:00	0.00%	0 (0.00%)	$0.00 (0.00%)
9. ad-review-tool.twitter.biz / referral	**1** **(0.18%)**	100.00%	1 (0.19%)	100.00%	1.00	00:00:00			
10. adwords.google.com / referral	**1** **(0.18%)**	100.00%	1 (0.19%)	100.00%	1.00	00:00:00			



how did this end up in my news feed?

this was not possible 20 years ago.

- why?
- what did people do instead?

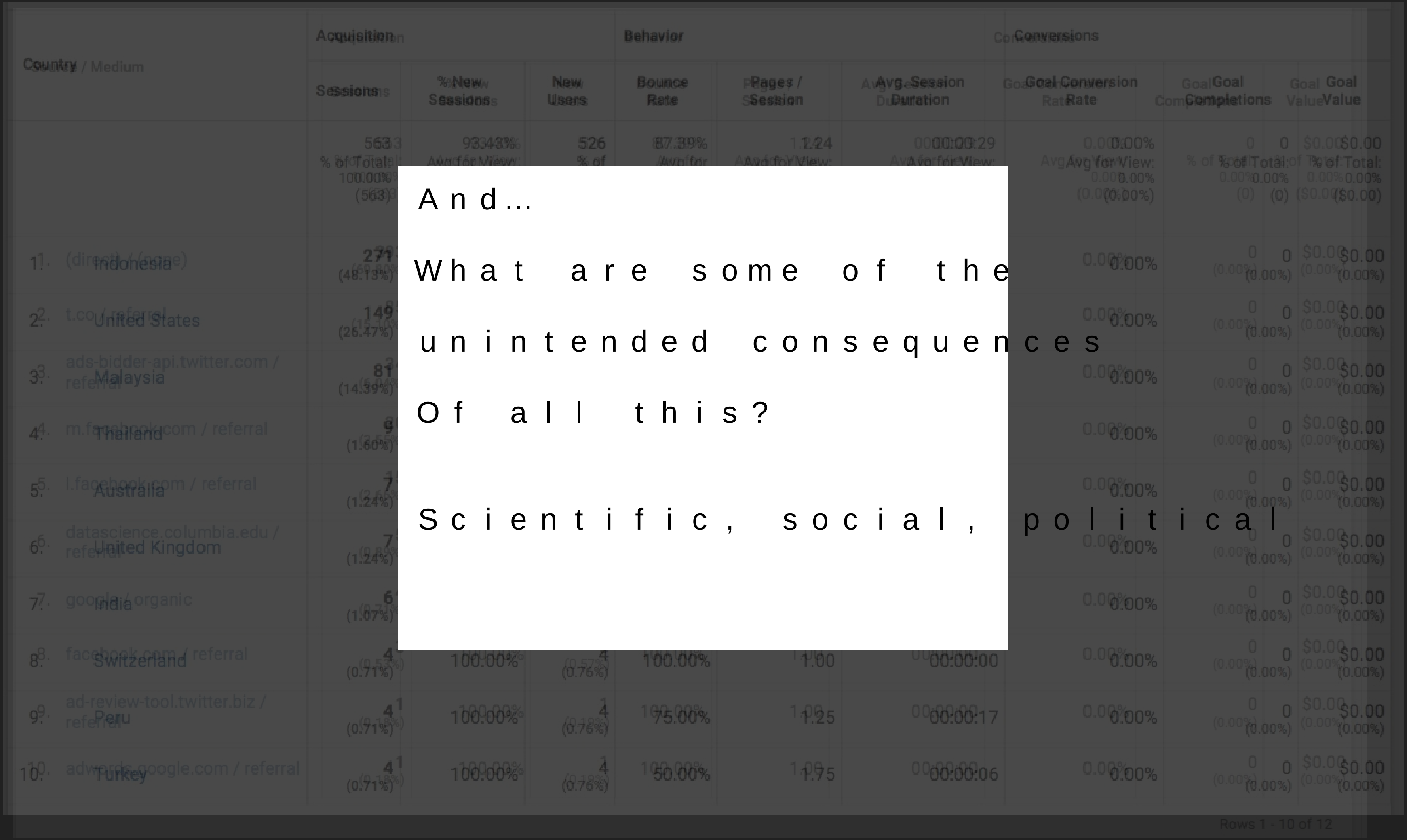


how did this end up in my news feed?

- math
- hardware
- system
- funding
- market
- regulation
- data

this was not possible 20 years ago.

- why?
- what did people do instead?



And…

What are some of the unintended consequences Of all this?

Scientific, social, political

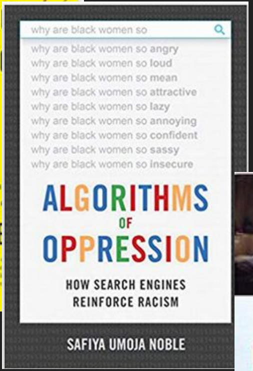
2017-09-05: cathy o'neil

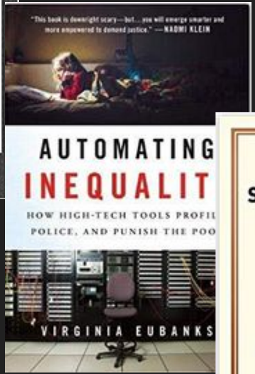
2018-01-08: safiya nob

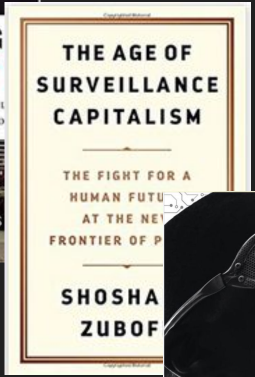
2018-01-23: virginia

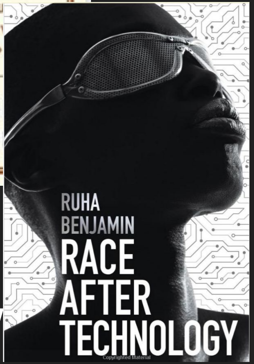
2019-01-15: shosh

2019-06-17: ru









something is w on the internet

Artificial intelligence (AI)

New AI can guess whether you're gay or straight from a photograph

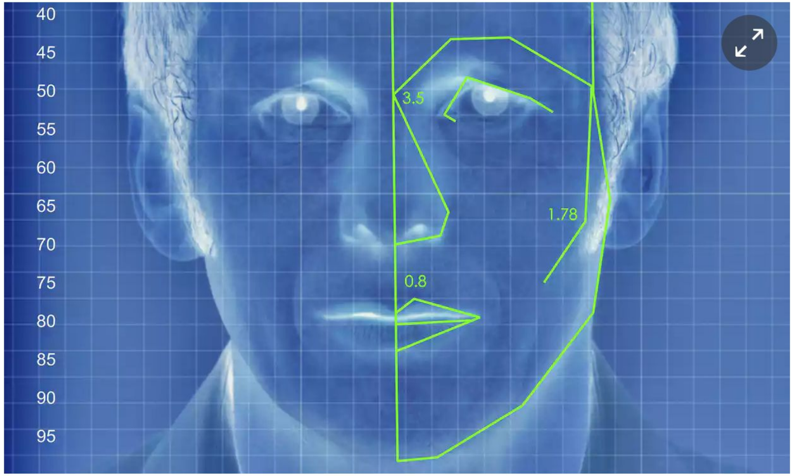
An algorithm deduced the sexuality of people on a dating site with up to 91% accuracy, raising tricky ethical questions

Sam Levin in San Francisco

@SamTLevin

Email

Fri 8 Sep 2017 00.46 BST



An illustrated depiction of facial analysis technology similar to that used in the experiment. Illustration: Alamy

Artificial intelligence can accurately guess whether people are gay or straight based on photos of their faces, according to new research that suggests machines can have significantly better "gaydar" than humans.

Public | 8

Deep neural networks are more accurate than humans at detecting sexual orientation from facial images.

Contributors: Yilun Wang, Michal Kosinski

Date created: 2017-02-15 08:37 AM | Last Updated: 2017-10-16 09:17 AM

Category: Project

Description: We show that faces contain much more information about sexual orientation than can be perceived and interpreted by the human brain. We used deep neural networks to extract features from 35,326 facial images. These features were entered into a logistic regression aimed at classifying sexual orientation. Given a single facial image, a classifier could correctly distinguish between gay and heterosexual men in 81% of cases, and in 74% of cases for women. Human judges achieved much lower accuracy: 61% for men and 54% for women. The accuracy of the algorithm increased to 91% and 83%, respectively, given five facial images per person. Facial features employed by the classifier included both fixed (e.g., nose shape) and transient facial features (e.g., grooming style). Consistent with the prenatal hormone theory of sexual orientation, gay men and women tended to have gender-atypical facial morphology, expression, and grooming styles. Prediction models aimed at gender alone allowed for detecting gay males with 57% accuracy and gay females with 58% accuracy. Those findings advance our understanding of the origins of sexual orientation and the limits of human perception. Additionally, given that companies and governments are increasingly using computer vision algorithms to detect people's intimate traits, our findings expose a threat to the privacy and safety of gay men and women.

The New York Times

The Secretive Company That Might End Privacy as We Know It

A little-known start-up helps law enforcement match photos of unknown people to their online images — and "might lead to a dystopian future or something," a backer says.

NY , 2020-01-18

Bernard Parker, left, was rated

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

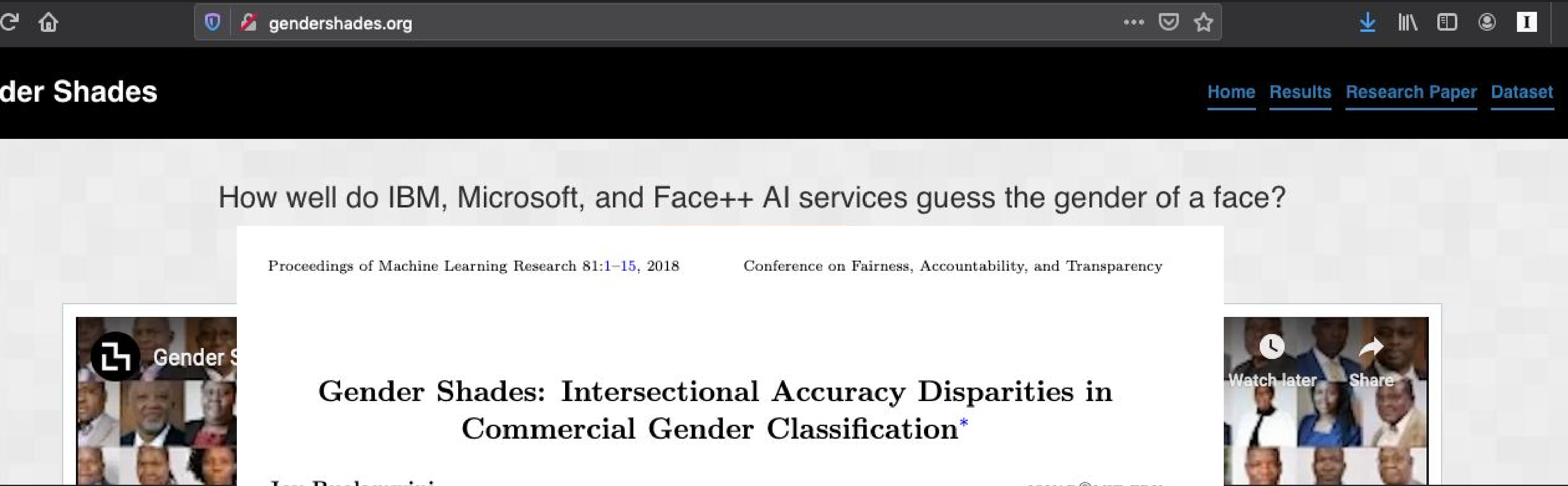
by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

ON A SPRING AFTERNOON IN 2014, Brisha Borden was running late to pick up her god-sister from school when she spotted an unlocked kid's blue Huffy bicycle and a silver Razor scooter. Borden and a friend grabbed the bike and scooter and tried to ride them down the street in the Fort Lauderdale suburb of Coral Springs.

Just as the 18-year-old girls were realizing they were too big for the tiny conveyances — which belonged to a 6-year-old boy — a woman came running after them saying, "That's my kid's stuff." Borden and her friend immediately dropped the bike and scooter and walked away.

But it was too late — a neighbor who witnessed the heist had already called the police. Borden and her friend were arrested and charged with burglary and petty theft for the items, which were valued at a total of $80.

Gendershades project: Buc



Proceedings of Machine Learning Research 81:1–15, 2018 Conference on Fairness, Accountability, and Transparency

Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*

Joy Buolamwini JOYAB@MIT.EDU
MIT Media Lab 75 Amherst St. Cambridge, MA 02139

Timnit Gebru TIMNIT.GEBRU@MICROSOFT.COM
Microsoft Research 641 Avenue of the Americas, New York, NY 10011

Editors: Sorelle A. Friedler and Christo Wilson

Automated Inference on Criminality using Face Images

Xiaolin Wu
McMaster University
Shanghai Jiao Tong University
xwu510@gmail.com

Xi Zhang
Shanghai Jiao Tong University
zhangxi_19930818@sjtu.edu.cn

Abstract

We study, for the first time, automated inference on criminality based solely on still face images, which is free of any biases of subjective judgments of human observers.

management science, criminology, etc.

In all cultures and all periods of recorded human history, people share the belief that the face alone suffices to reveal innate traits of a person. Aristotle in his famous work Prior Analytics asserted, "It is possible to infer character

“In all cultures
recorded human h
share the belief
suffices to reve
person.”

(a) Three samples in criminal ID photo set S_c.

(b) Three samples in non-criminal ID photo set S_n

Figure 1. Sample ID photos in our data set.

2. Data preparation

Still happening! Paper ye

Article | Open Access | Published: 11 January 2021

Facial recognition technology can expose political orientation from naturalistic facial images

Michal Kosinski

Scientific Reports **11**, Article number: 100 (2021) | Cite this article

Abstract

Ubiquitous facial recognition technology can expose individuals' political orientation, as faces of liberals and conservatives consistently differ. A facial recognition algorithm was applied to naturalistic images of 1,085,795 individuals to predict their political orientation by comparing

NEWS

Historians Politely Remind Nation To Check What's Happened In Past Before Making Any Big Decisions

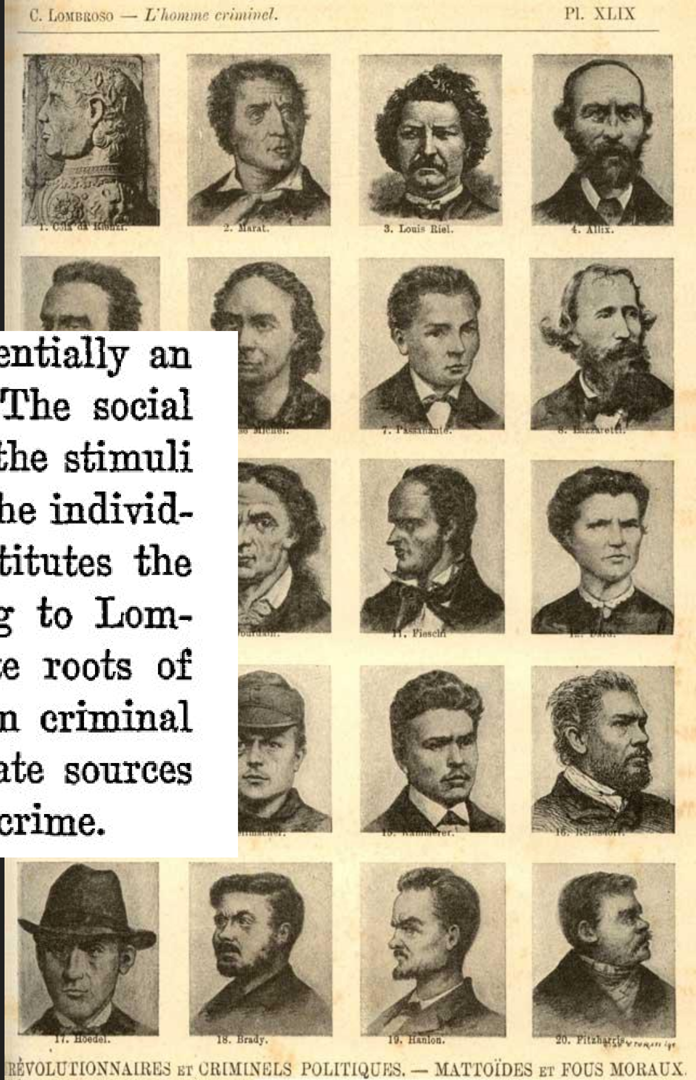
9/28/11 9:00am • SEE MORE: SCIENCE & TECHNOLOGY

t h e   o n i o n   d o t   c o m

Trying to avoid repeating bad things we did in the past is a good idea, historians say.

We've been here

Lombroso believed, in other words, that the criminal was essentially an organic anomaly, partly pathological and partly atavistic. The social causes of crime were at most, according to Lombroso, simply the stimuli which called forth the organic and psychical abnormalities of the individual. While the removal of the social causes of crime constitutes the immediate practical problem before criminologists, according to Lombroso, because they are the exciting causes, yet the ultimate roots of crime lie in the atavistic and degenerate heredity of the born criminal and the criminaloid, and only the extirpation of these ultimate sources of criminality can afford a final solution of the problem of crime.



We've been here before

Medium

Sign in

Blaise Aguera y Arcas Follow

Blaise Aguera y Arcas leads Google's AI group in Seattle. He founded Seadragon, and was one of the creators of Photosynth at Microsoft.

May 6, 2017 · 38 min read

Physiognomy's New Clothes

by Blaise Agüera y Arcas, Margaret Mitchell and Alexander Todorov

We've been here before

Medium

Sign in

Blaise Aguera y Arcas Follow

Blaise Aguera y Arcas leads Google's AI group in Seattle. He founded Seadragon, and was one of the creators of Photosynth at Microsoft.

Jan 11 · 15 min read

Do algorithms reveal sexual orientation or just expose our stereotypes?

by Blaise Agüera y Arcas, Alexander Todorov and Margaret Mitchell

We've been here before

Medium

Sign in

This doesn't negate the privacy concerns the authors and various commentators have raised, but it emphasizes that such concerns relate less to AI per se than to mass surveillance, which is troubling regardless of the technologies used (even when, as in the days of the Stasi in East Germany, these were nothing but paper files and audiotapes). Like computers or the internal combustion engine, AI is a general-purpose technology that can be used to automate a great many tasks, including ones that should not be undertaken in the first place.

We are hopeful about the confluence of new, powerful AI technologies with social science, but not because we believe in reviving the 19th century research program of inferring people's inner character from their outer appearance. Rather, we believe AI is an essential tool for understanding

Statistical sciences alwa

The aspiration for a "science"

Central to development of

Statistics

And the

Data sciences

(from 1800s to present day!

Statistical sciences alwa

Dream of sciences of social c

Central to development of

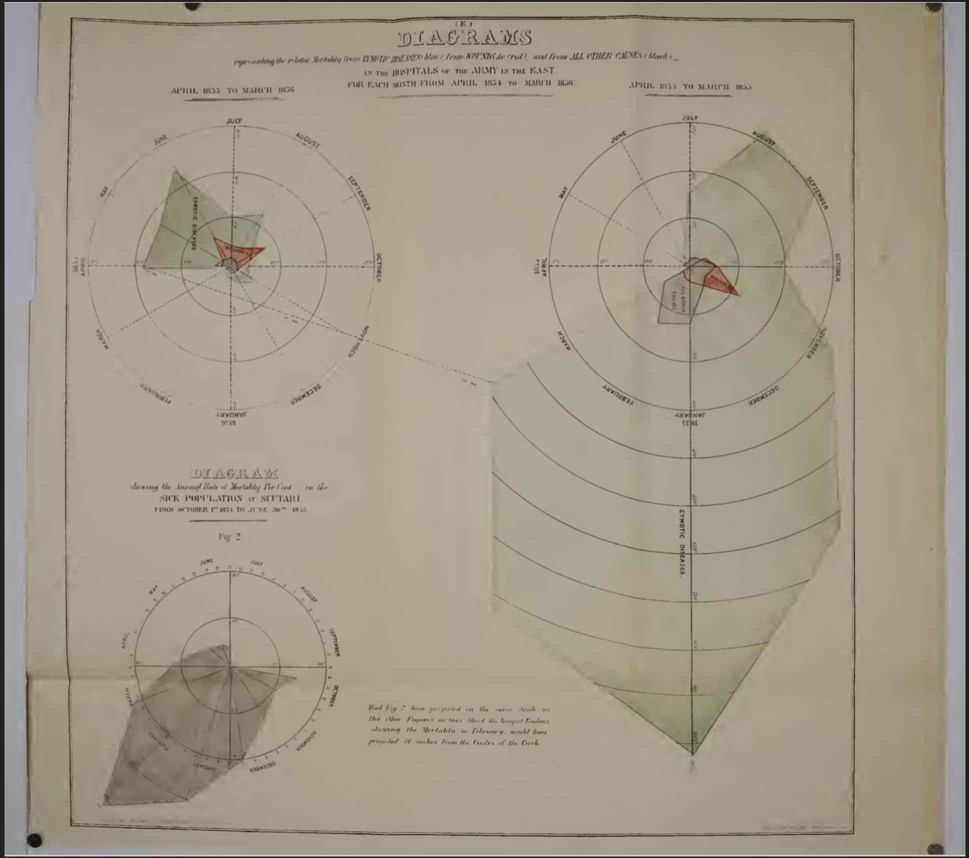
Statistics

And the

Data sciences

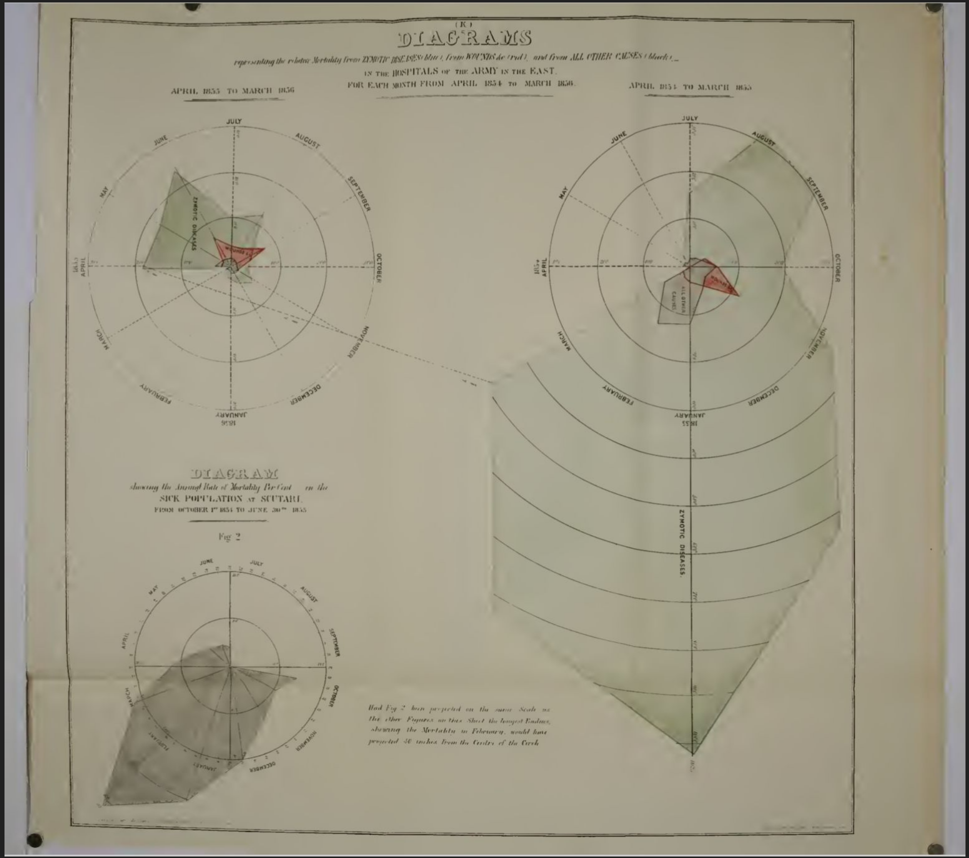
Florence Nightingale & Da

"Experience has s
special informati
application of th
science in preser
drain on our home
exhaust our means
introduction, the
sanitary system i
is of essential i
public interests.



Florence Nightingale & Da

"Upon the British
integrity of that
moment appears to
conquering race m
possession."



Arc of class

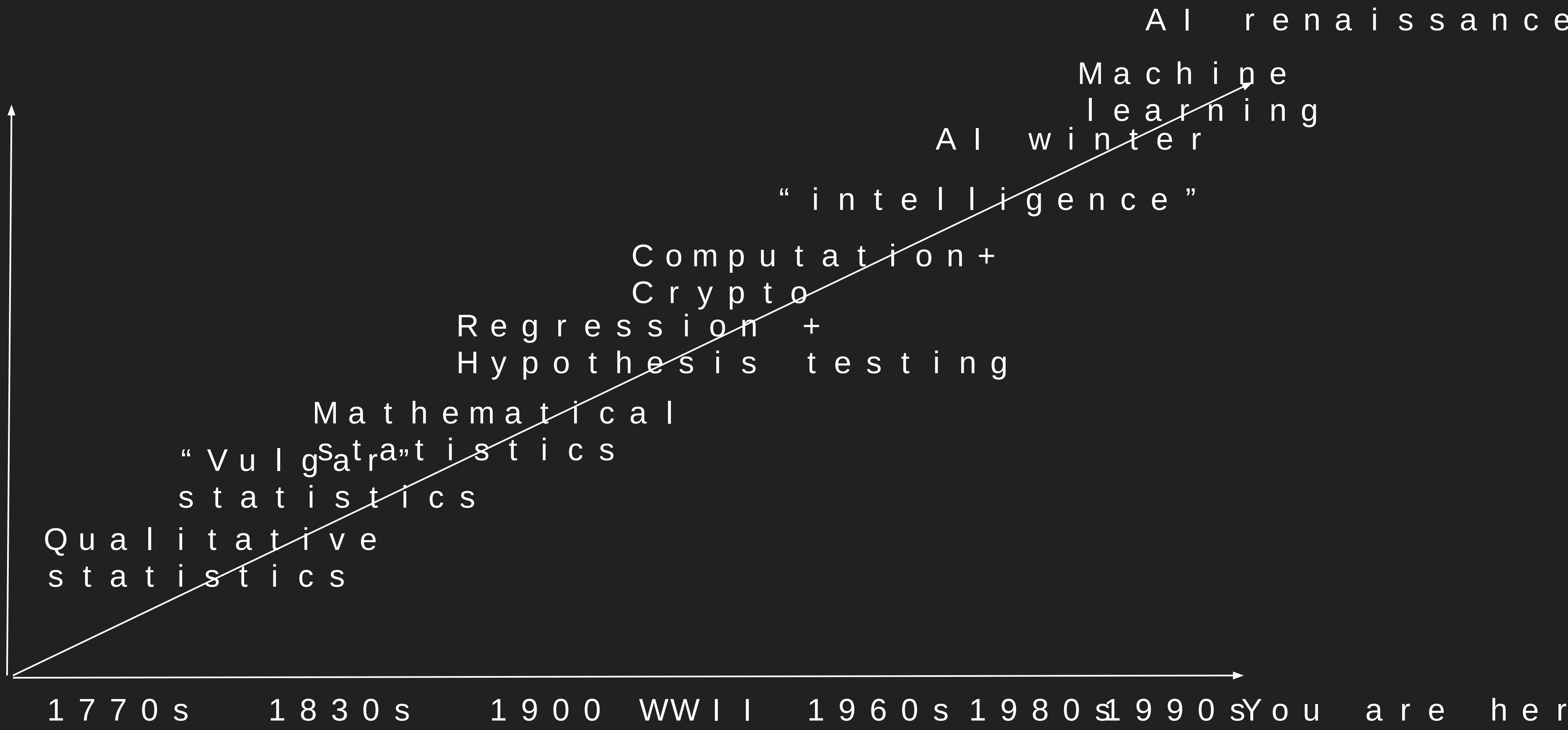
This class: three chronol

Data and Math

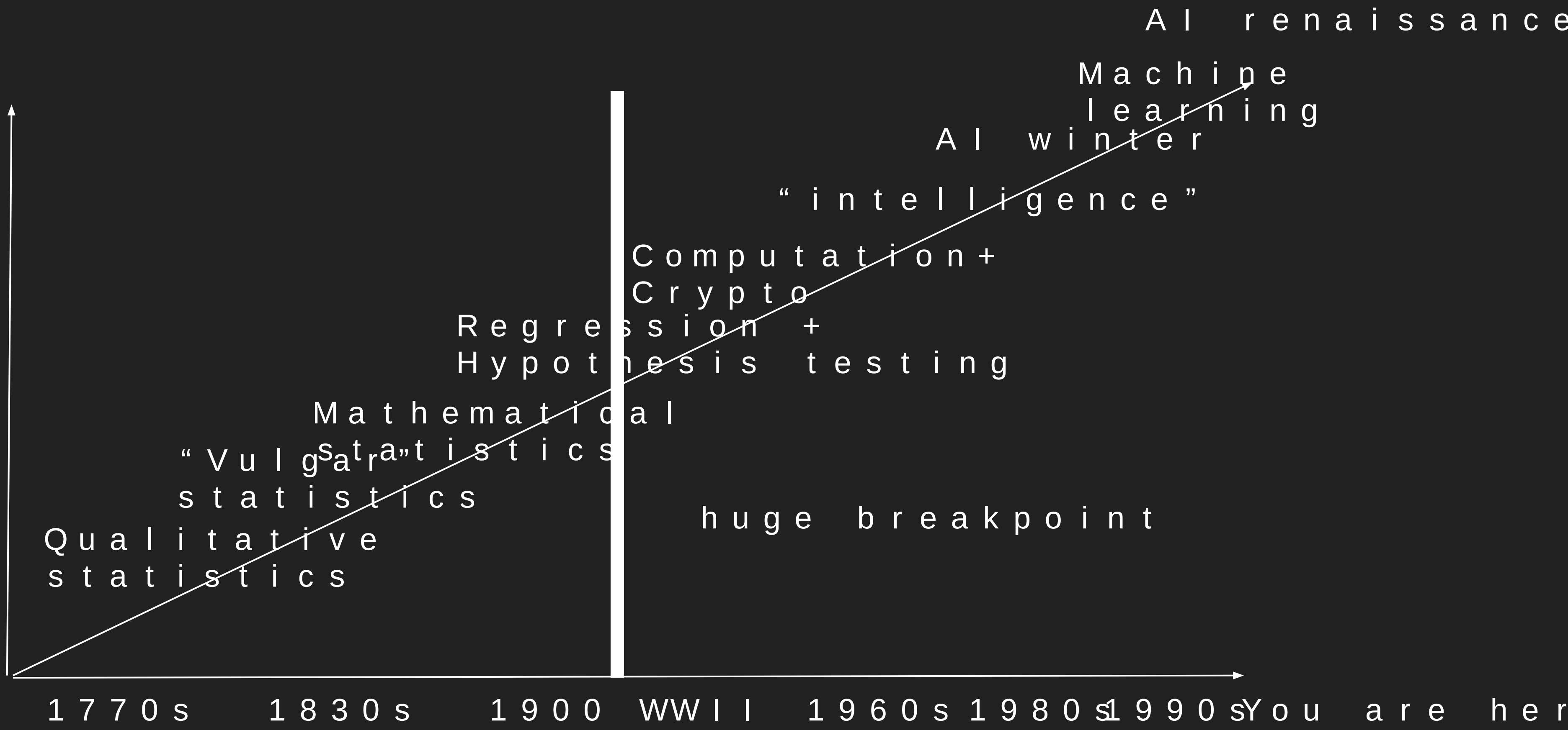
Data and Engineering

Data and Technology

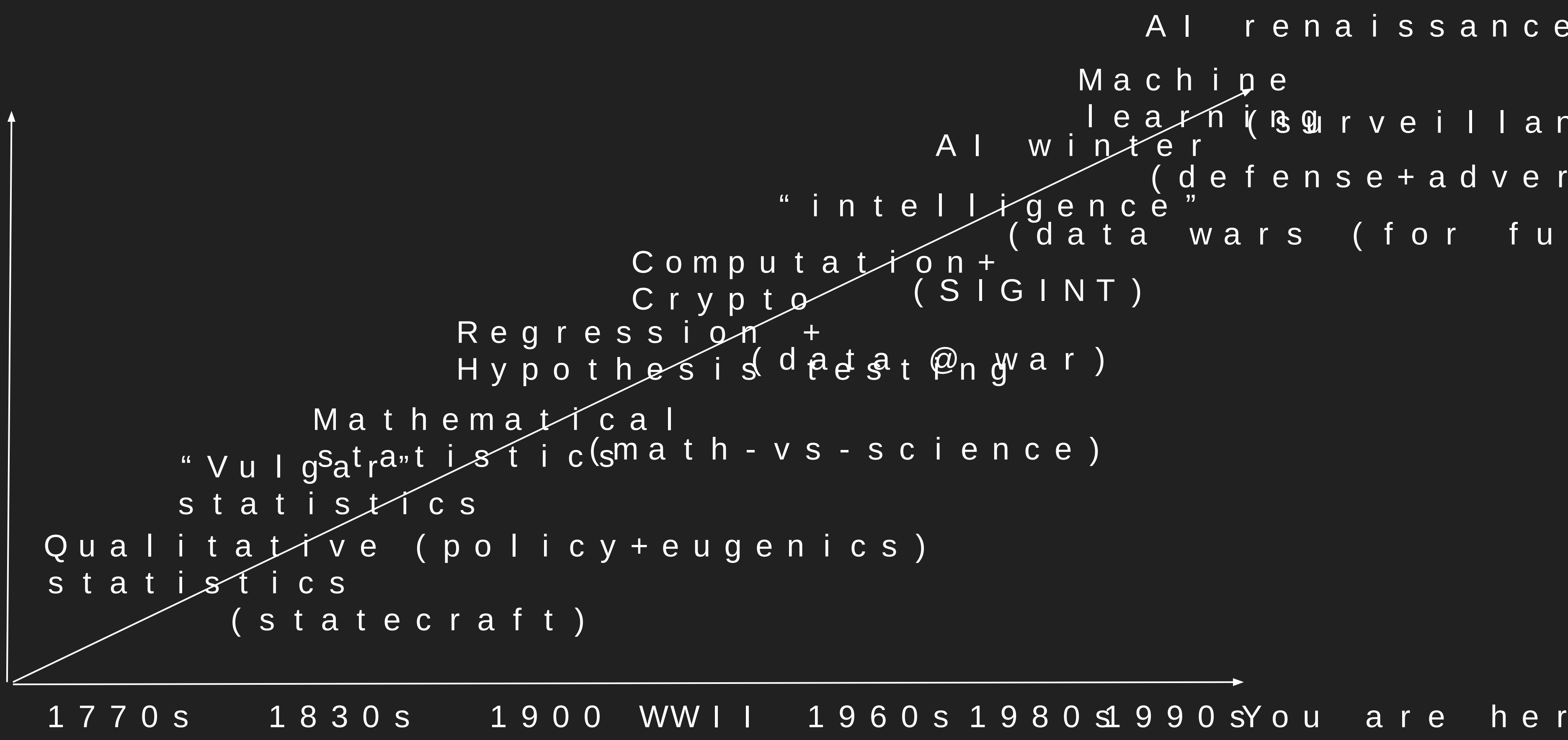
data 1770s-present: capabilit



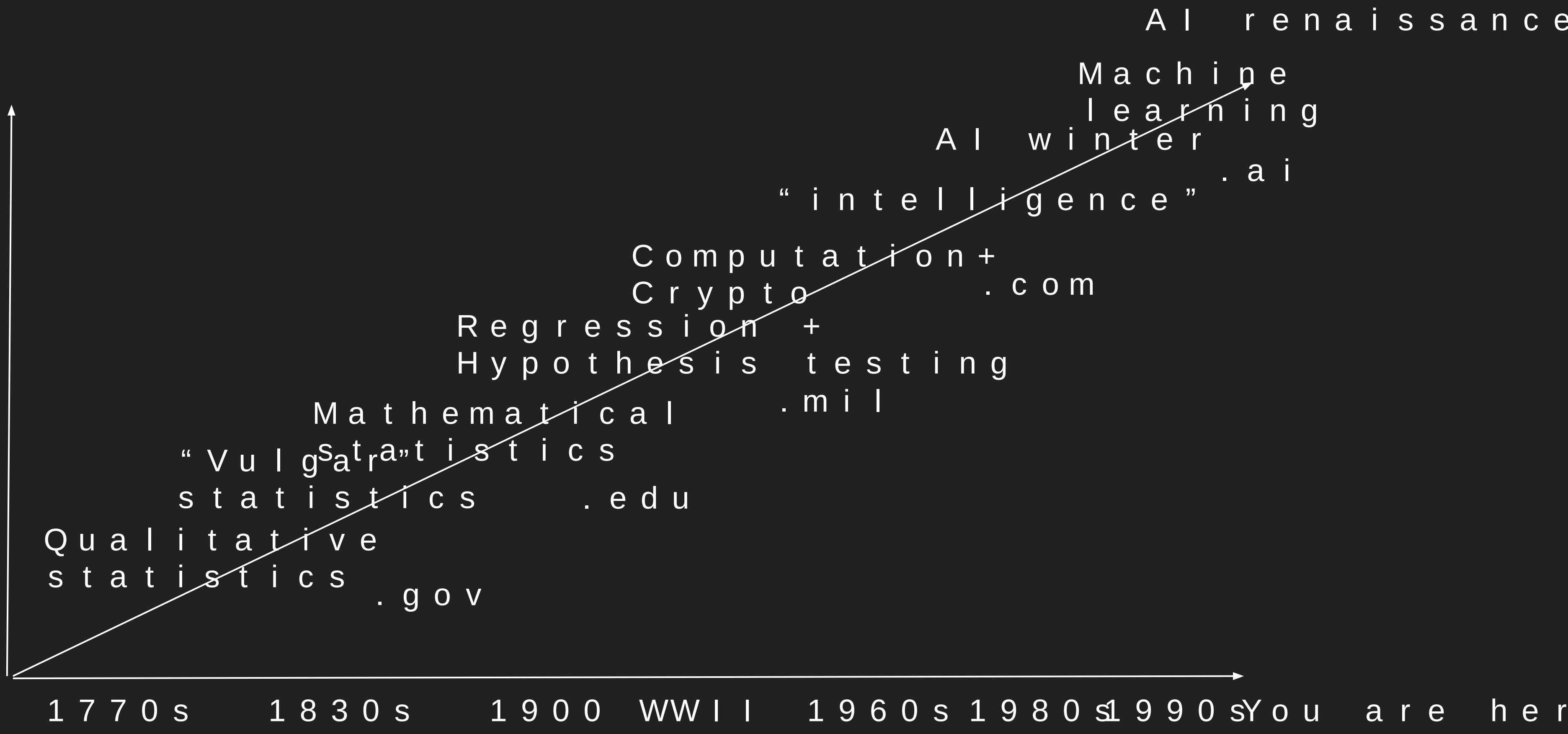
data 1770s-present: capabilit



data 1770s-present: capabilit



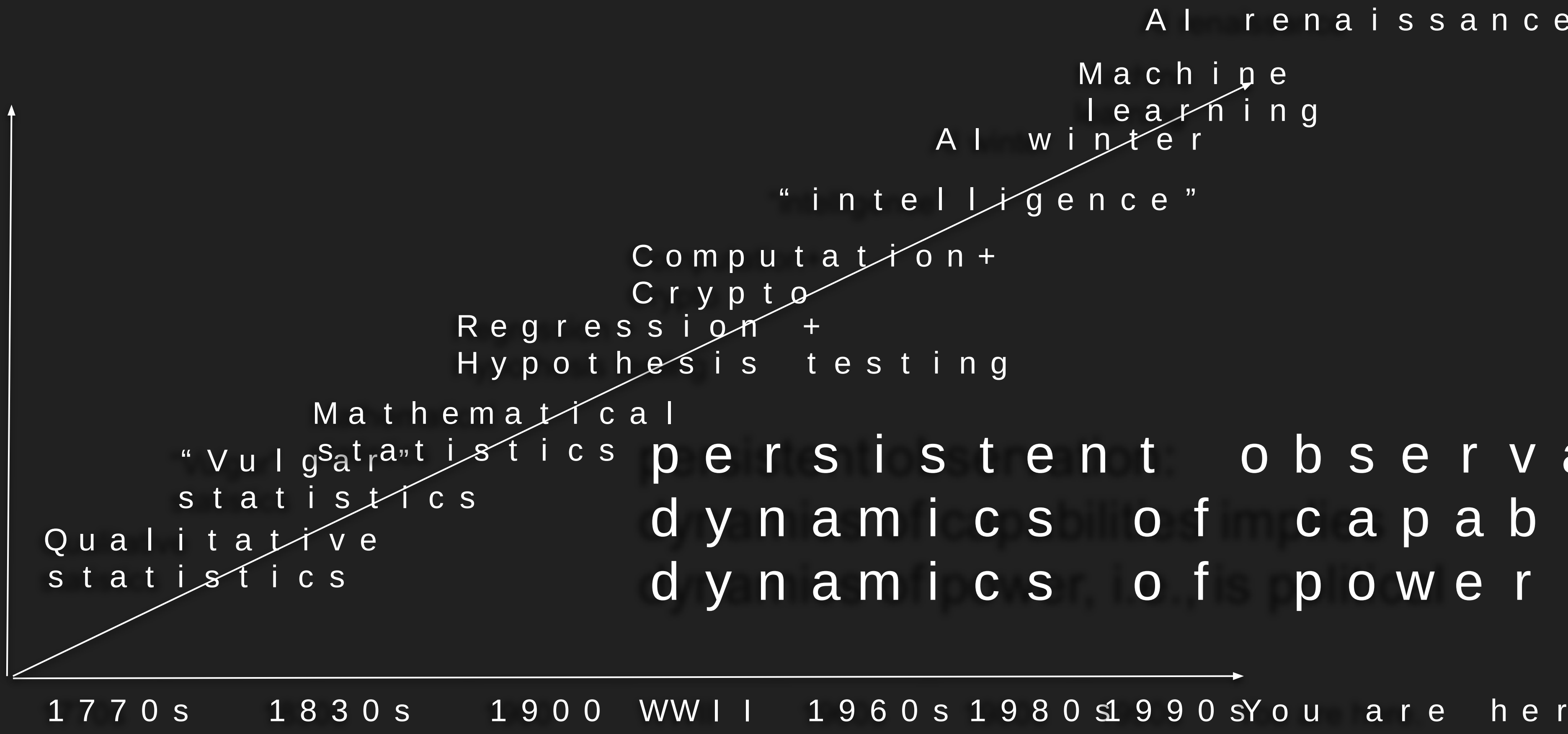
data 1770s-present: capabilit



data 1770s-present: capabilit

persistent observa
dynamics of capab
dynamics of power

data 1770s-present: capabilit

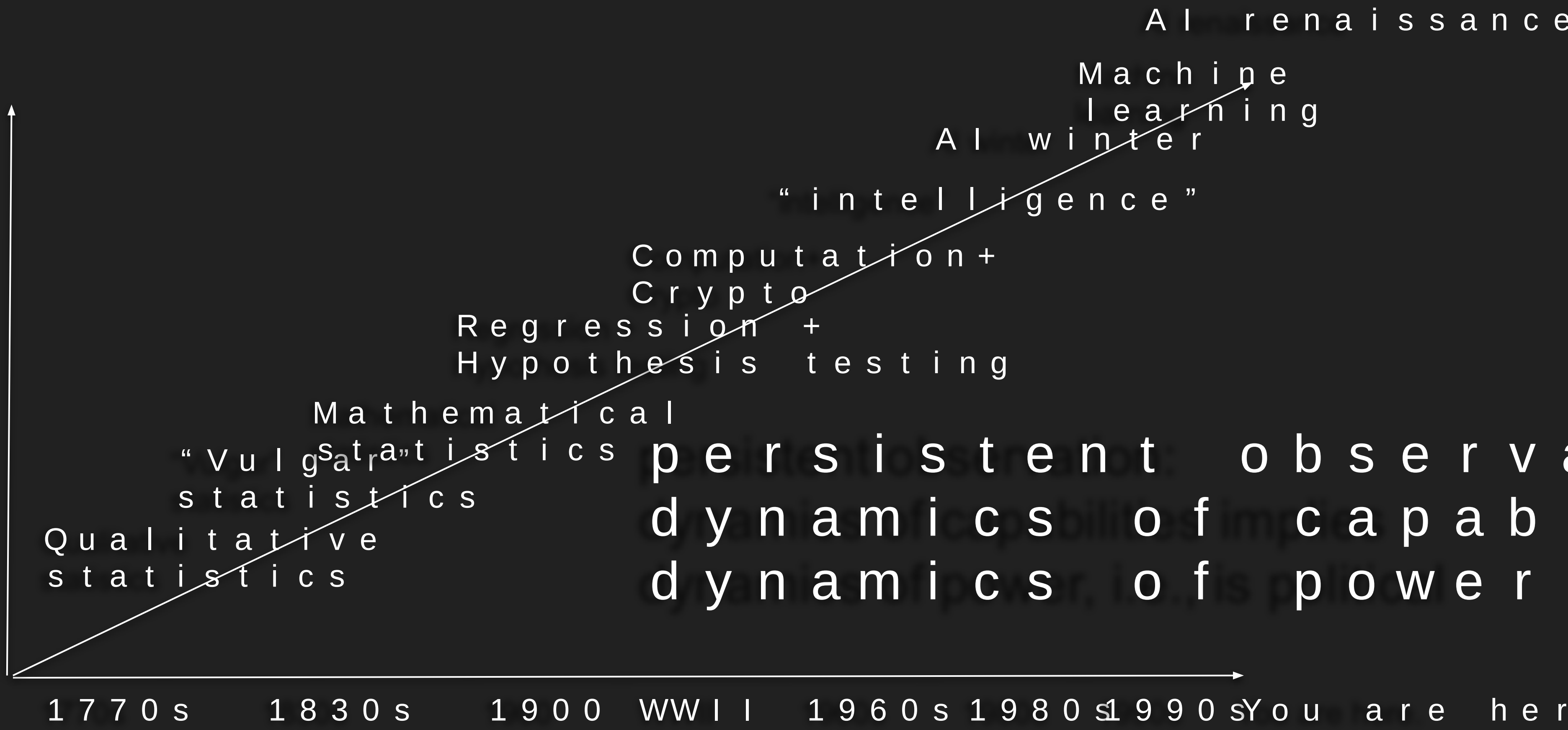


persistent observa
dynamics of capab
dynamics of power

Phillip Rogaway, "The Moral Work," December 2015

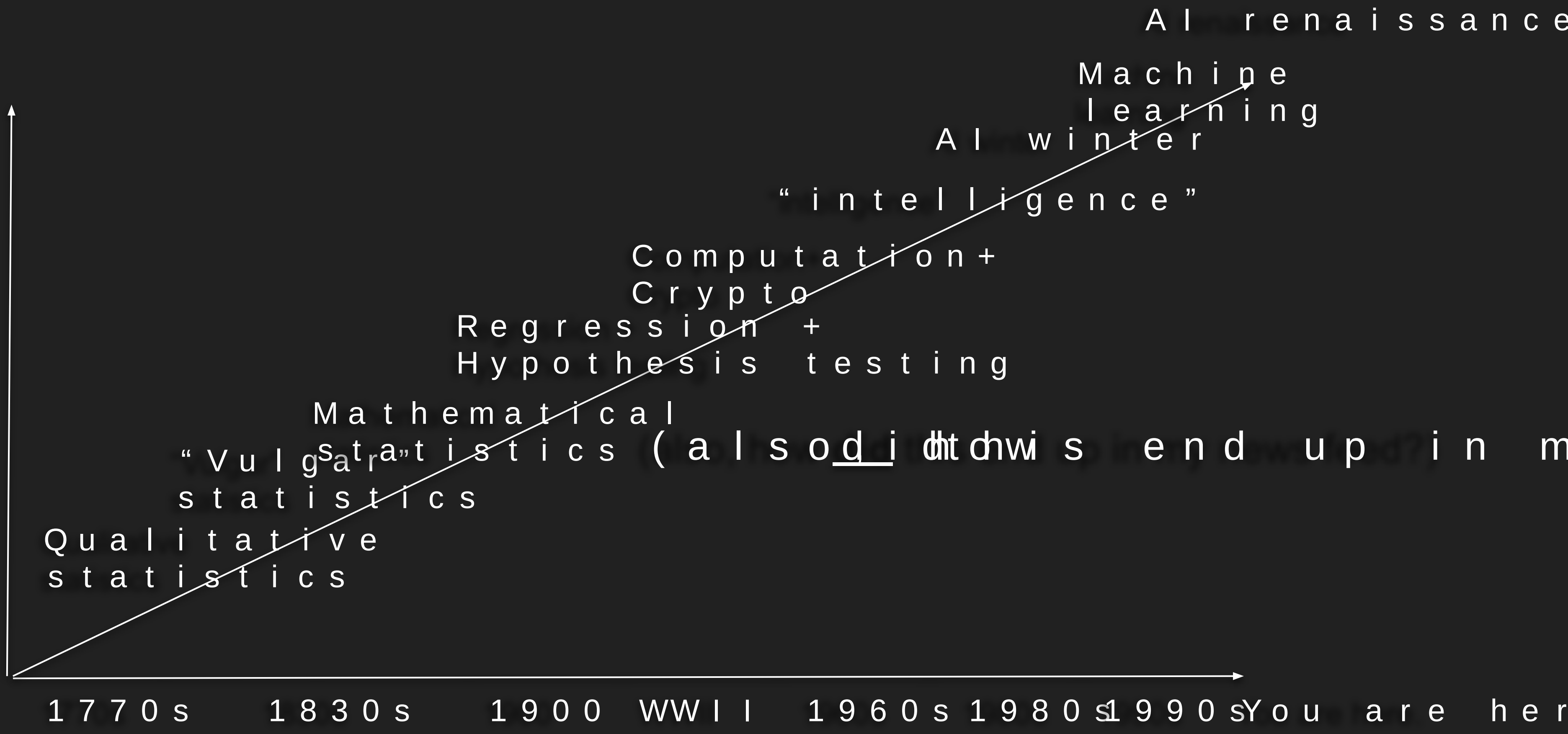
Abstract. Cryptography rearranges power: it configures who can do what, from what. This makes cryptography an inherently *political* tool, and it confers on the field an intrinsically *moral* dimension. The Snowden revelations motivate a reassessment of the political and moral positioning of cryptography. They lead one to ask if our inability to effectively address mass surveillance constitutes a failure of our field. I believe that it does. I call for a community-wide effort to develop more effective means

data 1770s-present: capabilit



persistent observa
dynamics of capab
dynamics of power

data 1770s-present: capabilit



How we know what the stat

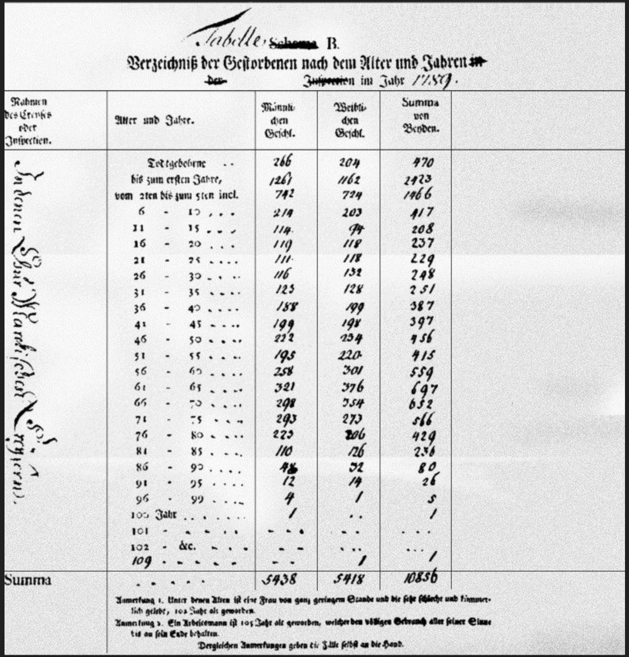


Tabelle ~~Schema~~ B.

Verzeichniß der Gestorbenen nach dem Alter und Jahren ~~in~~ ~~der~~ ~~Inspection~~ im Jahr 1789.

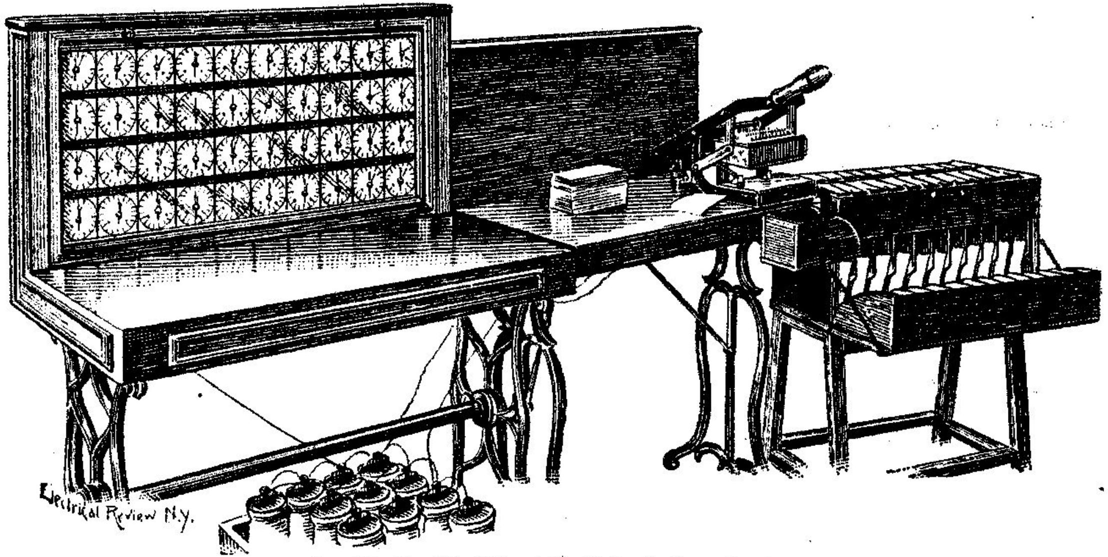
Nahmen des Creyses oder Inspection.	Alter und Jahre.	Männlichen Geschl.	Weiblichen Geschl.	Summa von Beyden.
[illegible]	Todtgebohrne . .	266	204	470
	bis zum ersten Jahre,	1261	1162	2423
	vom 2ten bis zum 5ten incl.	742	724	1466
	6 - 10 . . .	214	203	417
	11 - 15 . . .	114	94	208
	16 - 20 . . .	119	118	237
	21 - 25	111	118	229
	26 - 30	116	132	248
	31 - 35	123	128	251
	36 - 40	188	199	387
	41 - 45	199	198	397
	46 - 50	222	234	456
	51 - 55	195	220	415
	56 - 60	258	301	559
	61 - 65	321	376	697
	66 - 70	298	354	652
	71 - 75	293	273	566
	76 - 80	223	206	429
	81 - 85	110	126	236
	86 - 90	48	32	80
	91 - 95	12	14	26
	96 - 99	4	1	5
	100 Jahr	1	. .	1
	101 -	. . .	. . .	. . .
	102 - &c. . . .	. .	. . .	. . .
	109	. .	1	1
Summa	. . . . .	5438	5418	10856

Anmerkung 1. Unter denen Alten ist eine Frau von ganz geringem Stande und die sehr schlecht und kümmerlich gelebt, 102 Jahr alt geworden.

Anmerkung 2. Ein Arbeitsmann ist 105 Jahr alt geworden, welcher den völligen Gebrauch aller seiner Sinne bis an sein Ende behalten.

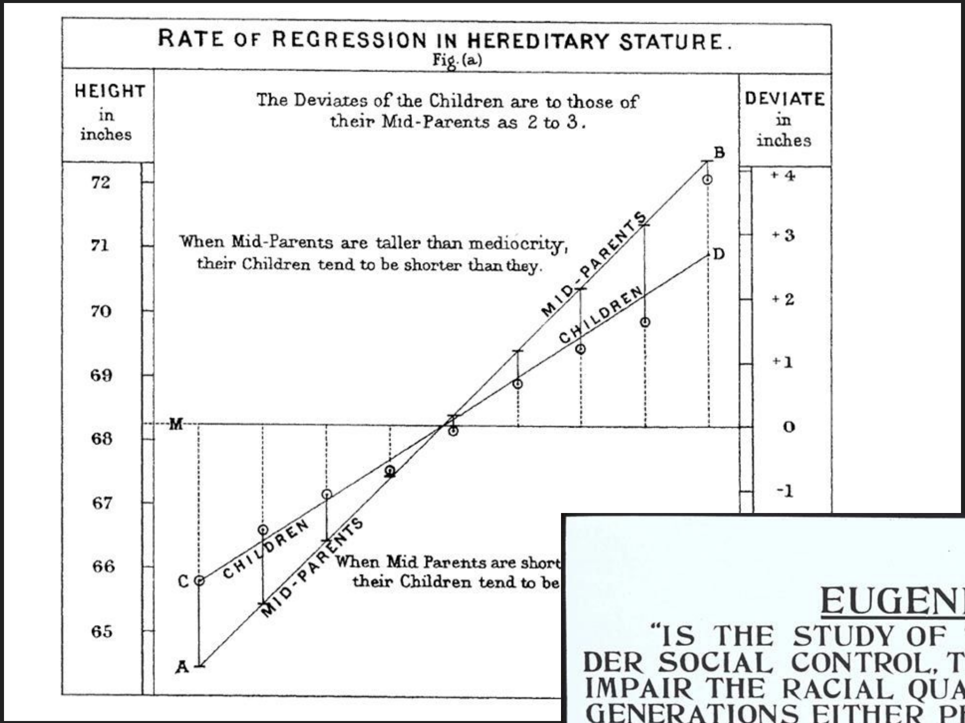
Dergleichen Anmerkungen geben die Fälle selbst an die Hand.

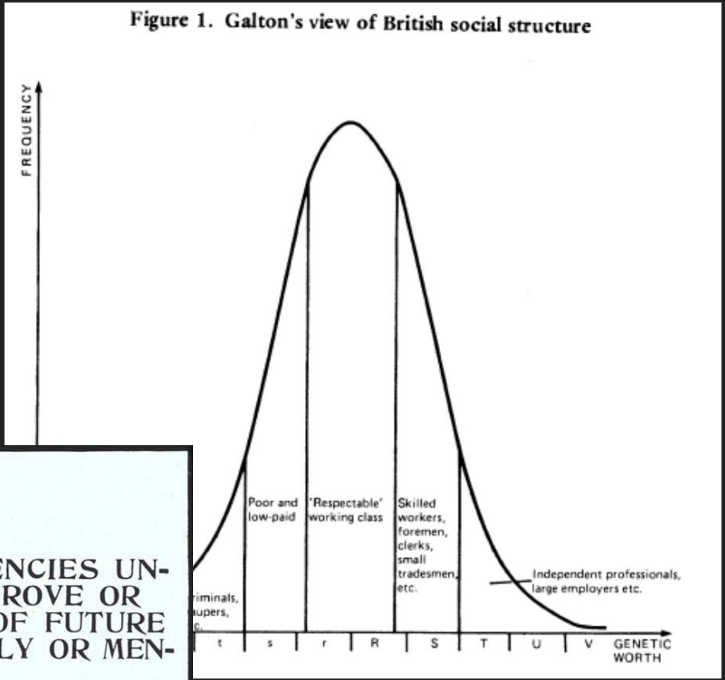
FIG. 2.

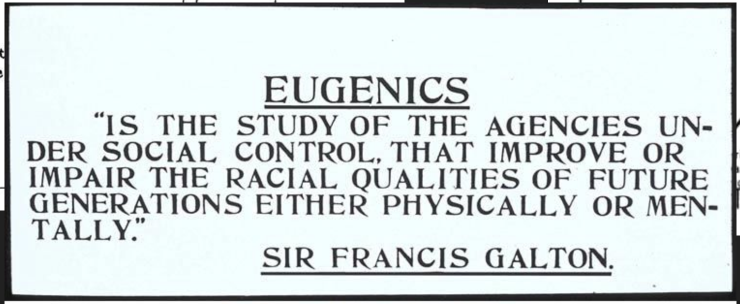


The Hollerith Electric Tabulating System.

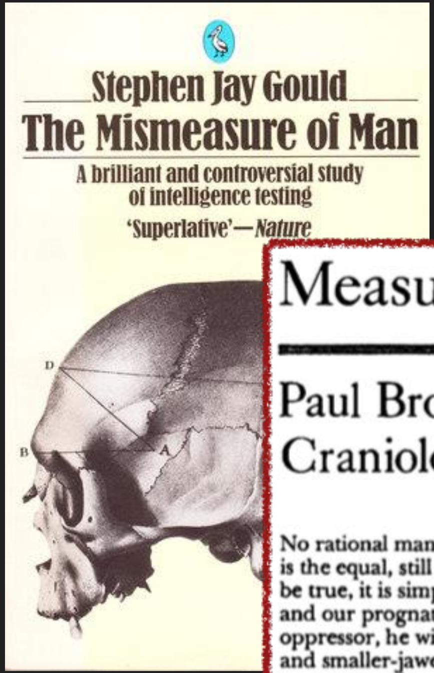
Regression, correlation a

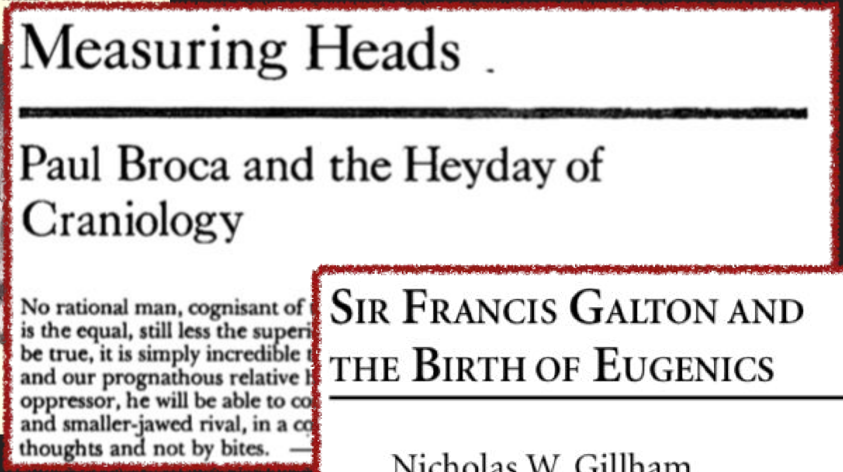






e.g., week 4 regression & quan

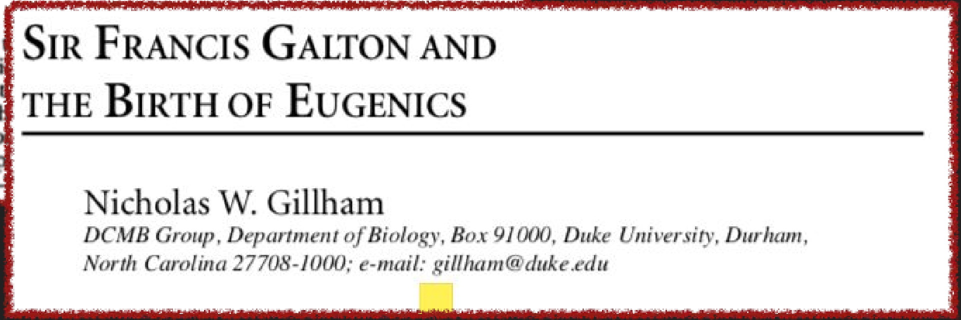




Measuring Heads

Paul Broca and the Heyday of Craniology

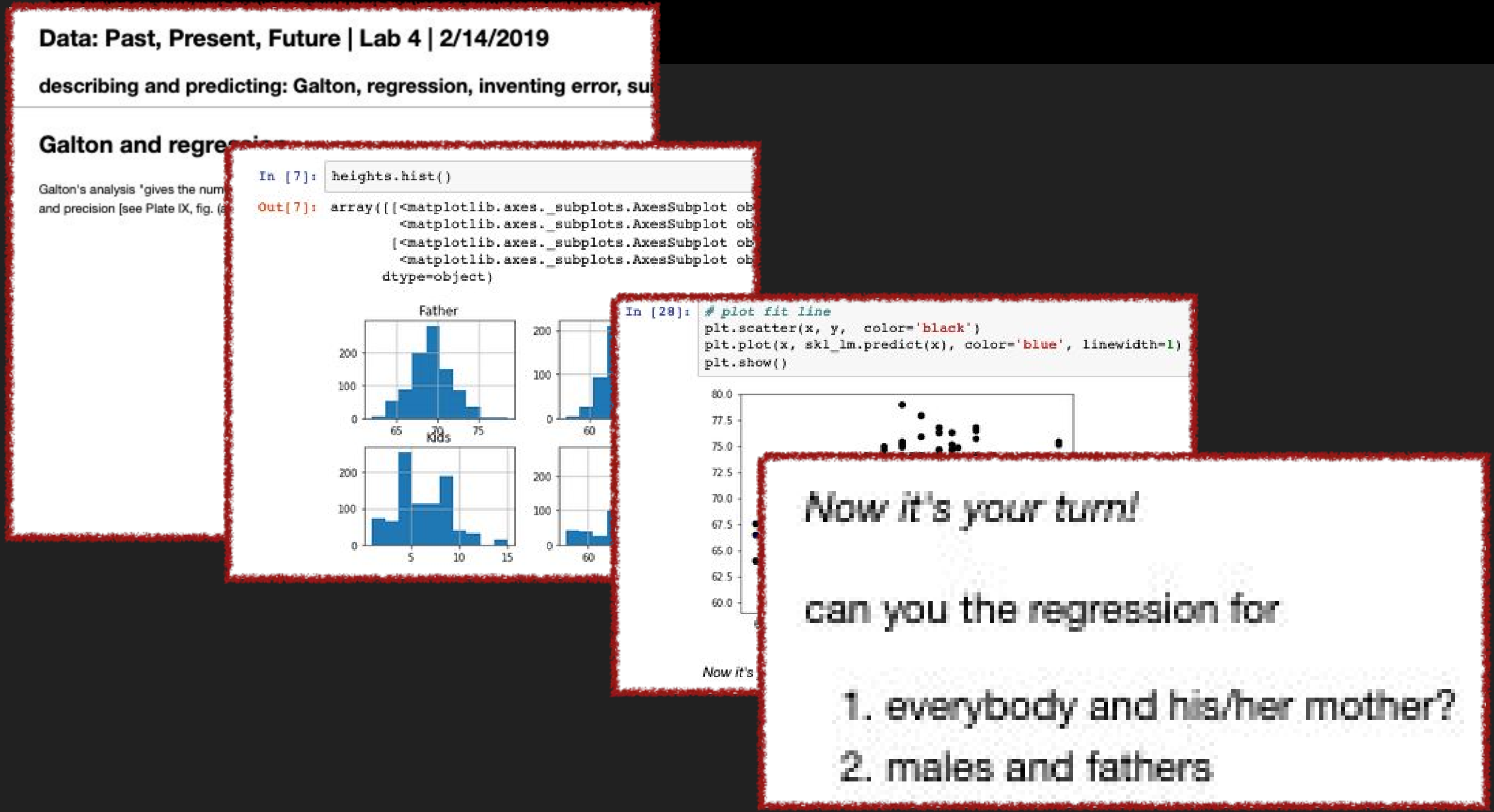
No rational man, cognisant of
is the equal, still less the superi
be true, it is simply incredible t
and our prognathous relative h
oppressor, he will be able to co
and smaller-jawed rival, in a co
thoughts and not by bites. —



SIR FRANCIS GALTON AND THE BIRTH OF EUGENICS

Nicholas W. Gillham
DCMB Group, Department of Biology, Box 91000, Duke University, Durham, North Carolina 27708-1000; e-mail: gillham@duke.edu

e.g., week 4 regression & quan



Experimental design, hypothes

To play this game with the greatest cha
success, the experimenter cannot affor
exclude the possibility of any possibl
arrangement of soil fertilities, and h
strategy is to equalize the chance tha
treatment shall fall on a de
it by chance himself.

-R.A. Fis

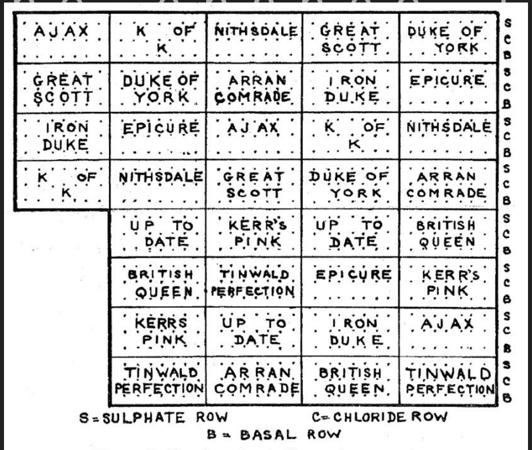
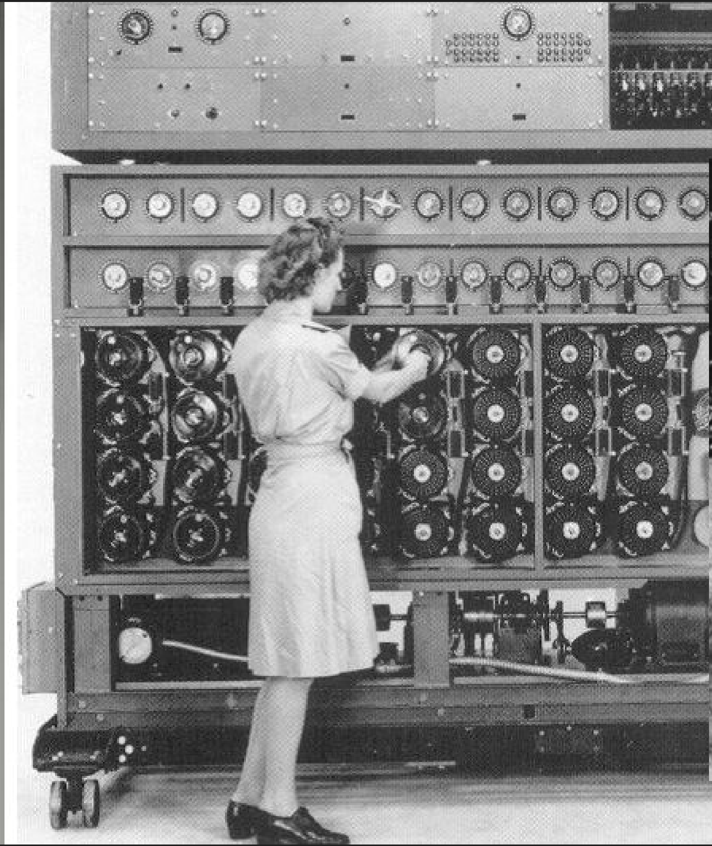
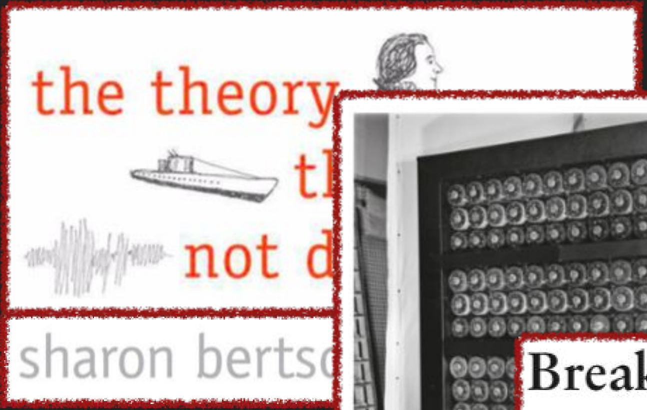


Diagram 1. Plan of experiment. Farmyard manure series.

World War 2: Turing and s



e.g., week 7 “women at the daw



Not a computer. The Bombe,
Alan Turing and Gordon We

Breaking Codes and Finding Trajectories: Women at the Dawn of the Digital Age

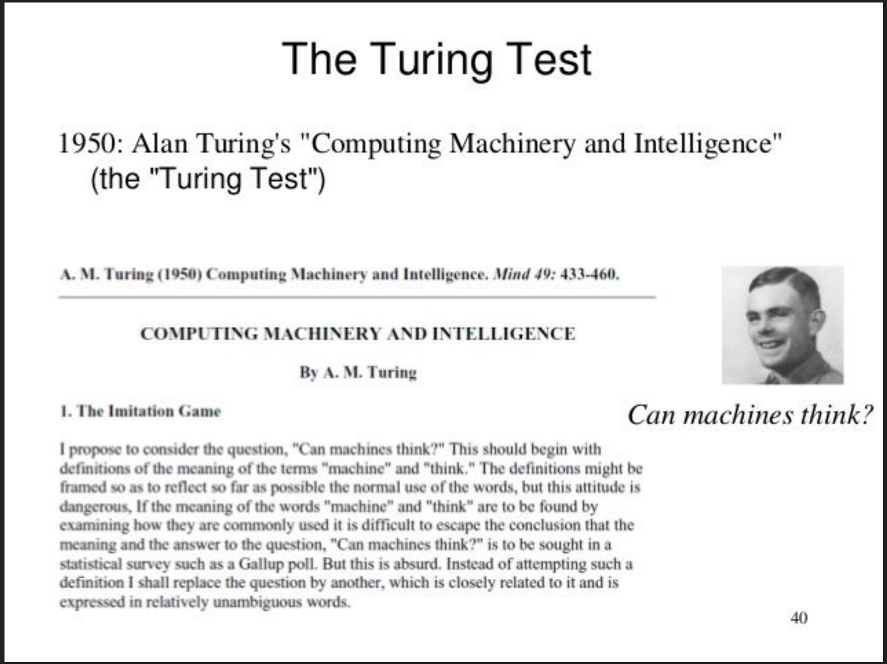
I was a Colossus operator, which we considered to be the crème de la crème. We felt we were “at the sharp end,” where there was a great tension and flow of adrenaline . . . operating those incredible machines.

—Jean Beech, Colossus operator[1]

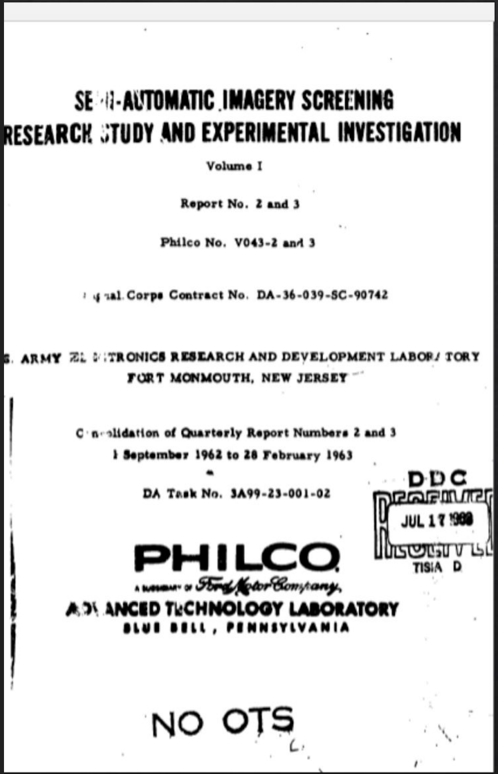
I don’t know if you can picture how exciting the ENIAC was to all of us. And we didn’t talk socially or any other time about anything else. It was—we discussed it almost all the time.

—Jean Jennings, ENIAC programmer[2]

AI and its many winters



Pattern Recognition to .



SE-AUTOMATIC IMAGERY SCREENING
RESEARCH STUDY AND EXPERIMENTAL INVESTIGATION

Volume I

Report No. 2 and 3

Philco No. V043-2 and 3

Signal Corps Contract No. DA-36-039-SC-90742

S. ARMY ELECTRONICS RESEARCH AND DEVELOPMENT LABORATORY
FORT MONMOUTH, NEW JERSEY

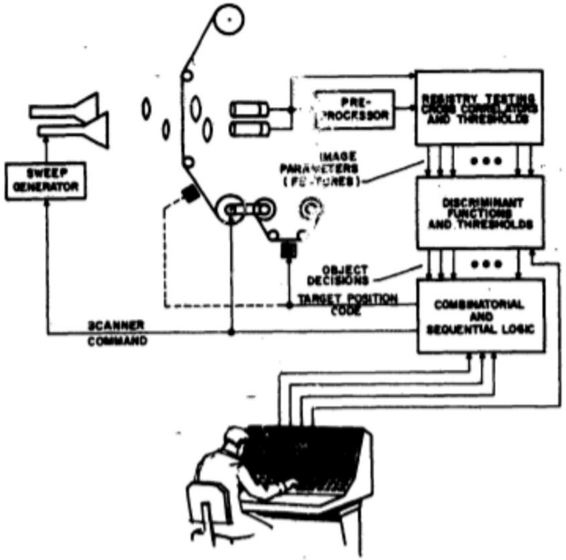
Consolidation of Quarterly Report Numbers 2 and 3
1 September 1962 to 28 February 1963

DA Task No. 3A99-23-001-02

DDC
JUL 17 1963
TISIA D

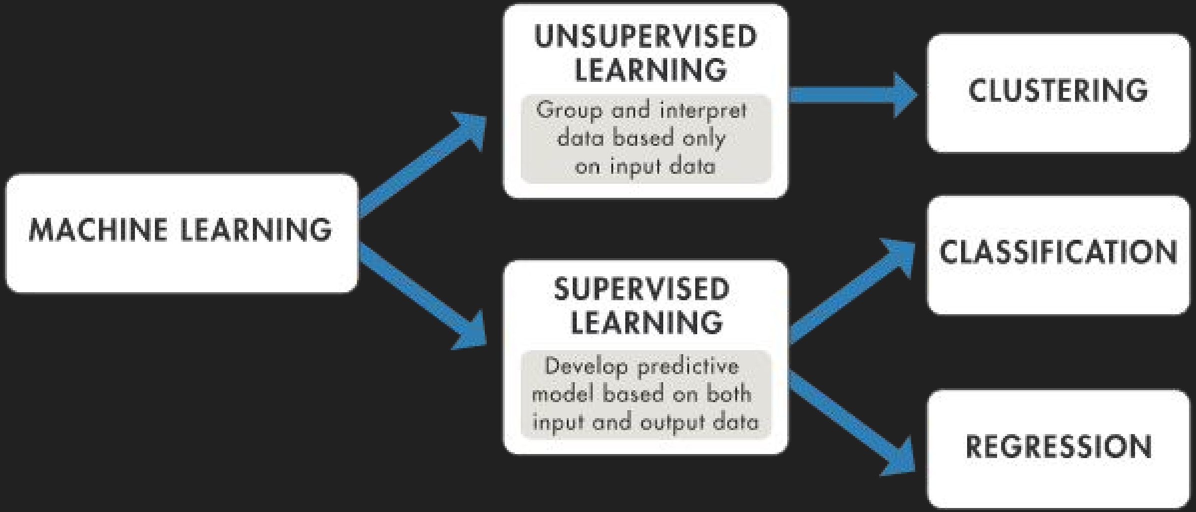
PHILCO
A SUBSIDIARY OF Ford Motor Company,
ADVANCED TECHNOLOGY LABORATORY
BLUE BELL, PENNSYLVANIA

NO OTS



Conceptual Block Diagram of Image Screening System

… Machine Learning

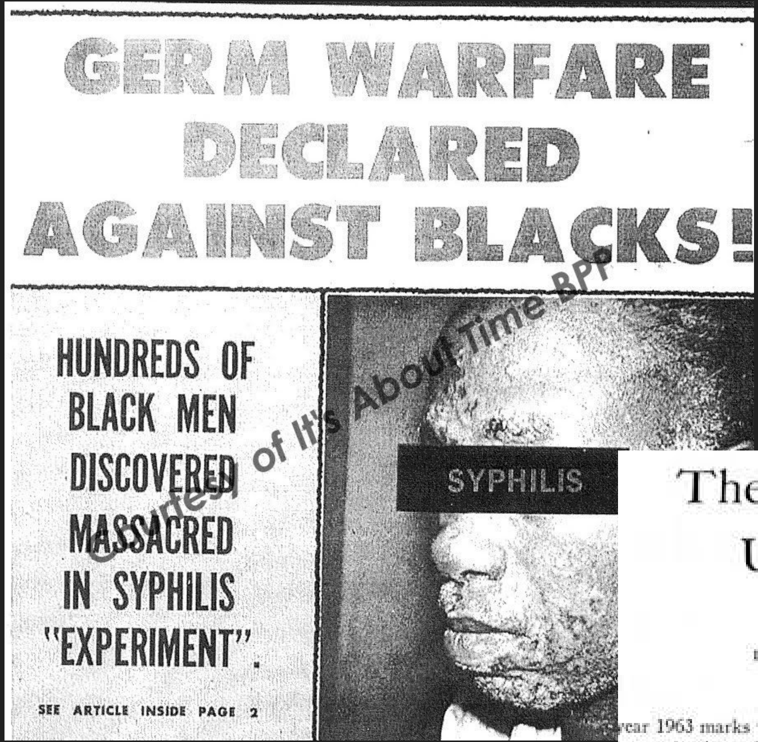


… and then to ethics & po

how did "ethics" meet "data

1. widely-condemned dumpste
2. defining "practical" eth
3. enforcing / designing fo
4. "AI ethics" esp. 2016 to

Research ethics (and the

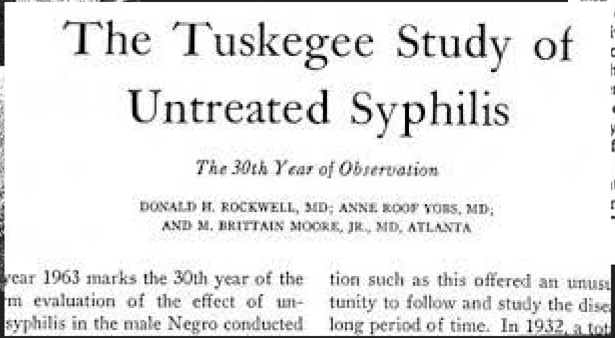


GERM WARFARE DECLARED AGAINST BLACKS!

HUNDREDS OF BLACK MEN DISCOVERED MASSACRED IN SYPHILIS "EXPERIMENT".

SEE ARTICLE INSIDE PAGE 2

SYPHILIS



The Tuskegee Study of Untreated Syphilis

The 30th Year of Observation

DONALD H. ROCKWELL, MD; ANNE ROOF YOBS, MD; AND M. BRITTAIN MOORE, JR., MD, ATLANTA

year 1963 marks the 30th year of the rm evaluation of the effect of un- syphilis in the male Negro conducted tion such as this offered an unusu tunity to follow and study the dise long period of time. In 1932, a tot

The New York Times

Syphilis Victims in U.S. Study Went Untreated for 40 Years

By JEAN HELLER
The Associated Press

WASHINGTON, July 25—For 40 years the United States Public Health Service has conducted a study in which human beings with syphilis, who were induced to serve as guinea pigs, have gone without medical treatment for the disease and a few have died of its effects, even though an ef- ive therapy was eventually overed.

he study was conducted to rmine from autopsies what disease does to the human y.

fficials of the health serv- who initiated the experi- t have long since retired. rent officials, who say they have serious doubts about the morality of the study, also say that it is too late to treat the syphilis in any surviving participants.

Doctors in the service say they are now rendering whatever other medical services they can give to the survivors while the study of the disease's effects continues.

Dr. Merlin K. DuVal, Assistant Secretary of Health, Education and Welfare for Health and Scientific Affairs, expressed shock on learning of the study. He said that he was making an immediate investigation.

The experiment, called the Tuskegee Study, began in 1932 with about 600 black men,

Research ethics (and the

Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer[a,1], Jamie E. Guillory[b,2], and Jeffrey T. Hancock[b,c]

[a]Core Data Science Team, Facebook, Inc., Menlo Park, CA 94025; and

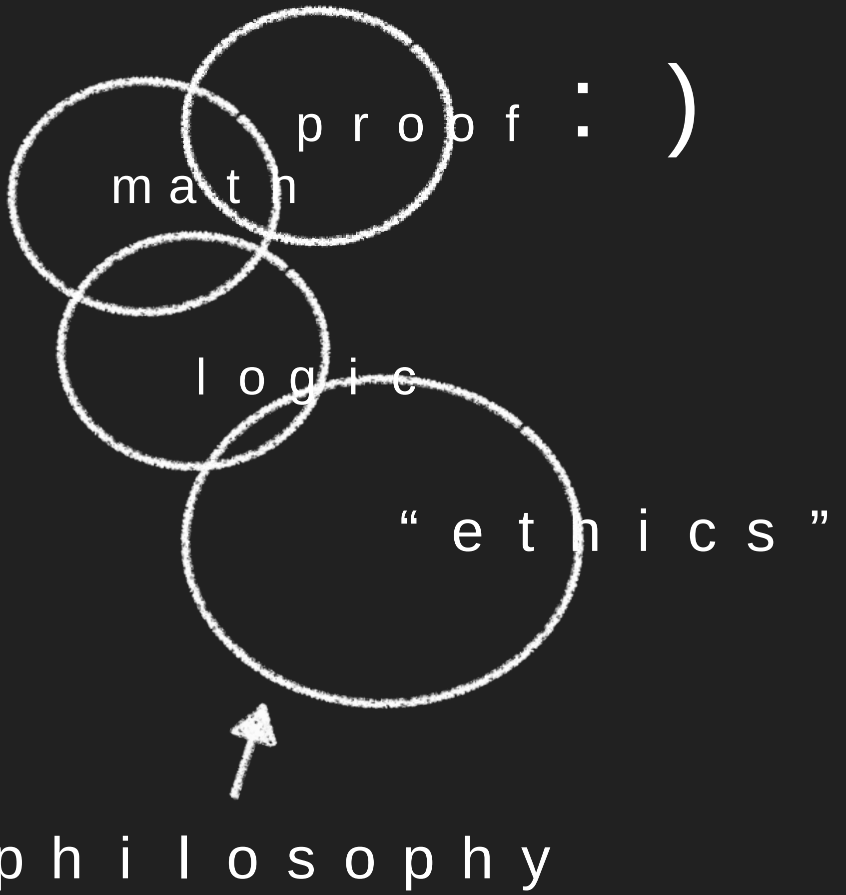
Departments of [b]Communication and

[c]Information Science, Cornell University, Ithaca, NY 14853

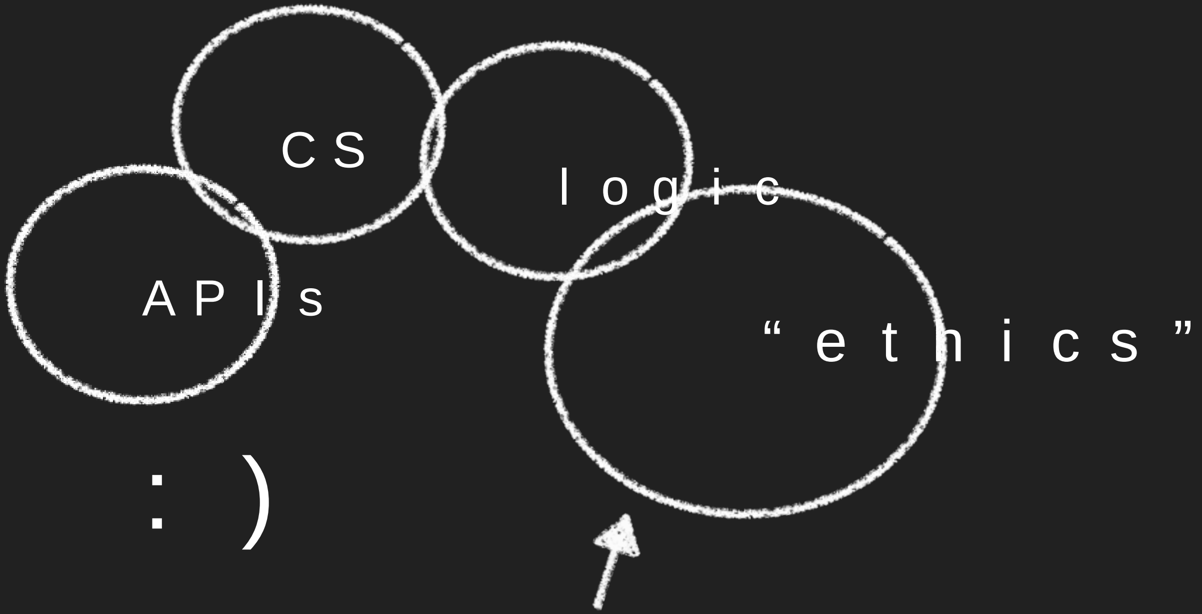
Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

Significance

We show, via a massive (*N* = 689,003) experiment on Facebook, that emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness. We provide experimental evidence that emotional contagion occurs without direct interaction between people (exposure to a friend expressing an emotion is sufficient), and in the complete absence of nonverbal cues.



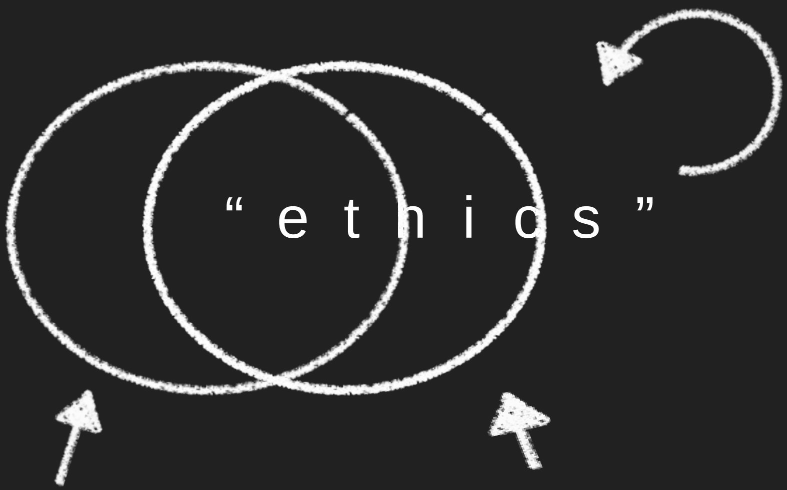
what we talk about when we talk



philosophy

what we talk about when we talk

what we talk about when we talk



philosophy sociology

(define) (design)

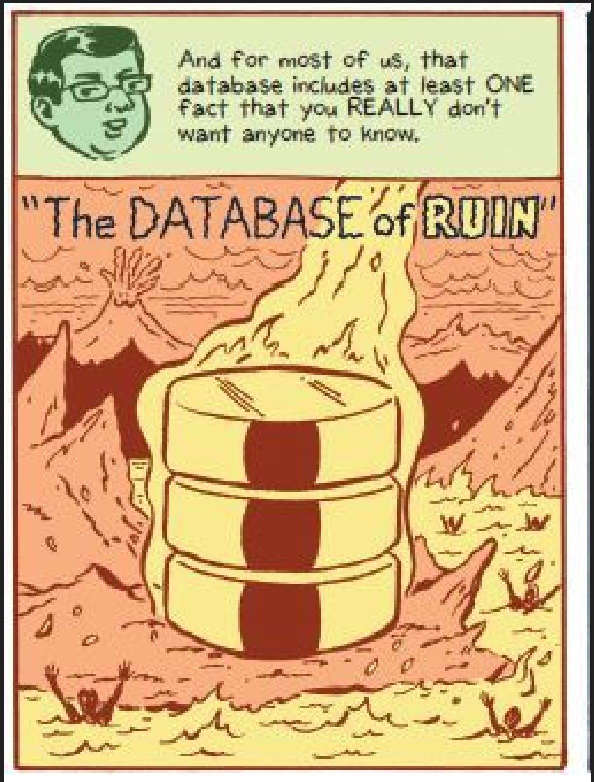
e.g., week 12 the ethics of da

Belmont principles (1978)

1. respect for personhood
   ~~- informed~~ consent -> autonomy
2. beneficence
   ~~- do no~~ harm -> balance risk+benefi
3. justice
   ~~- legal~~ -> fair, e.g., “veil of ig

gives analytical, hierarchical,
for ethical audit of decisions,
from which rules, code, “design

De-anonymization



Example of Re-identification

Public by Group Insurance Commission of Massachusetts

Voter register list of Massachusetts purchased with only 20$

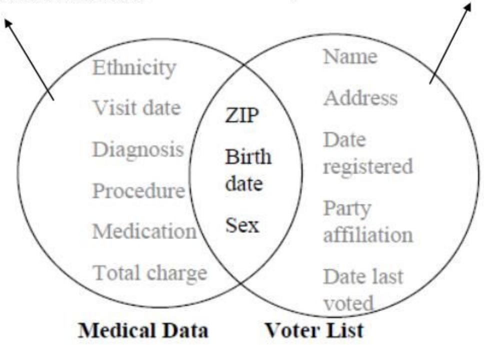


Figure 1 Linking to re-identify data

87% of Population in 1990. US are likely to be uniquely identified based on only on Zip, Birth and Sex

Sweeney, 2002

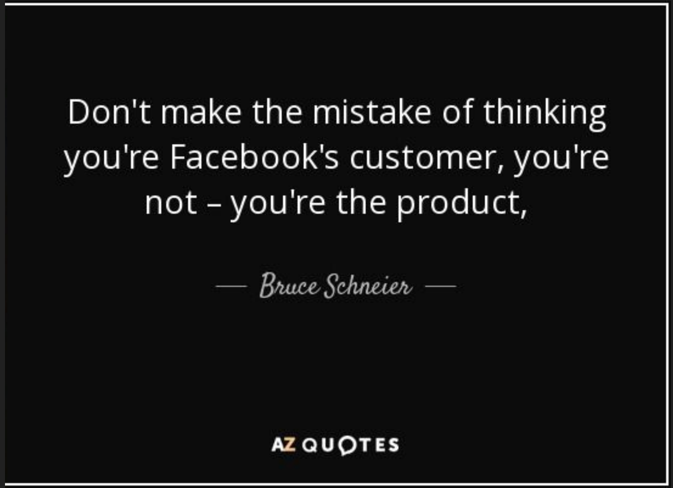
Socio-technological dynam

Census and governmen

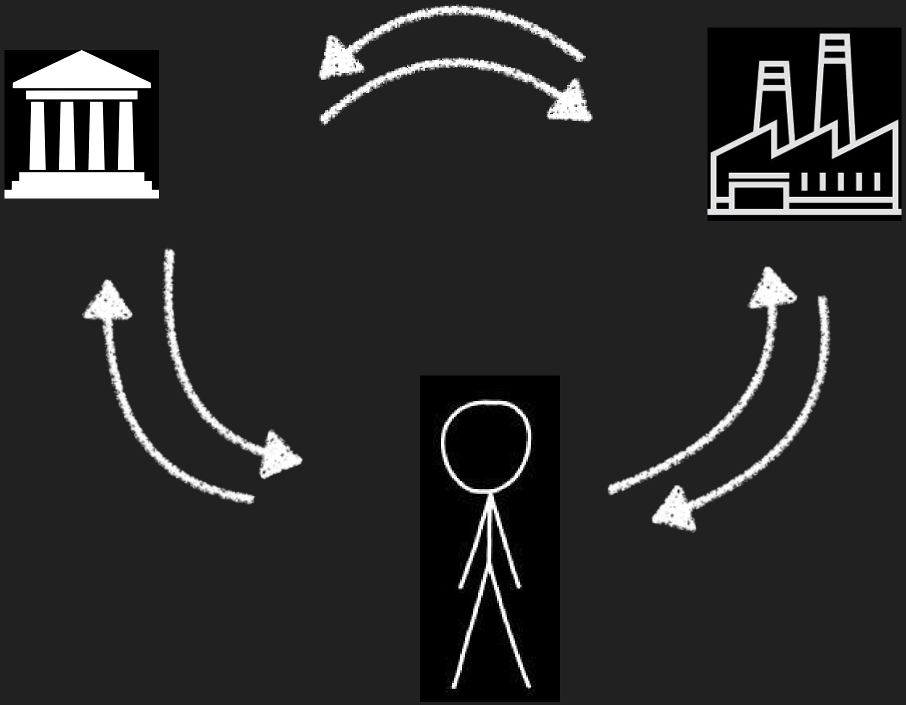
Information processi d digital computers

Always on network in

Silicon Valley and the At

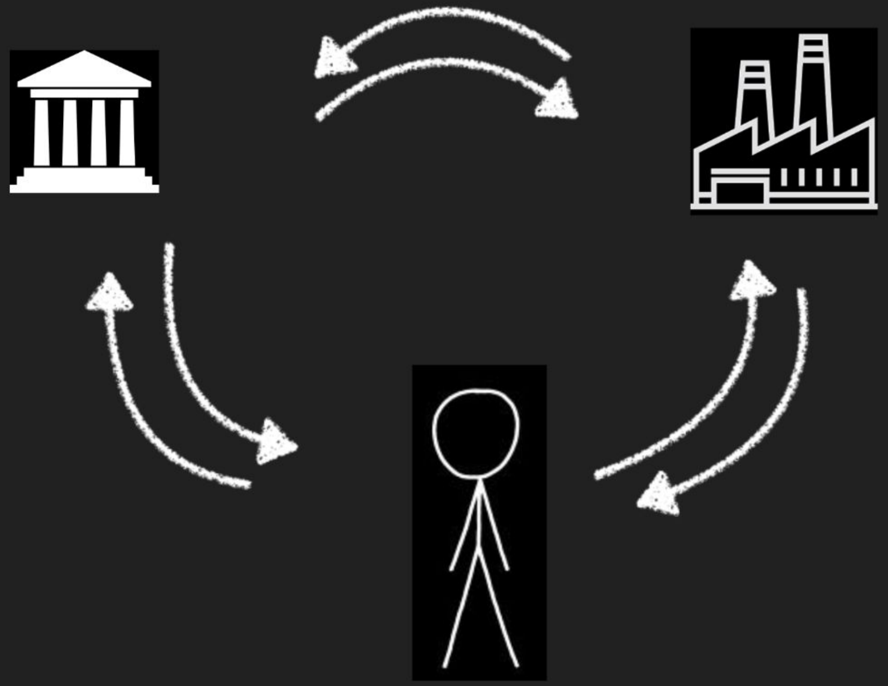


e.g., week 14 (future) solutio



3-player unstable game (adapted

e.g., week 14 (future) solutions



3-player unstable game (adapted from Janeway) u t i o


https://www.nytimes.com/2020/05/28/business/section-230-internet-speech.ht
by Lawmakers
Amazon Employees For Climate Justice @AMZNforClim... · Jan 9
Enough is enough. Amazon hosts Parler on @awscloud.
Hundreds of Twitter employees
https://www.nytimes.com/2021/01/04/technology/google-employees-union.html
The New York Times
Hundreds of Google Employees Unionize, Culminating Years of Activism
The creation of the union, a rarity in Silicon Valley, follows years of increasing outspokenness by Google workers. Executives have struggled to handle the change.
Like ou
Amazo
bullhorn to incite violence and attack our democracy.
buzzfeednews.com/article/johnpa... @JohnPaczkowski
Show this thread

o t e d

Tech, data, and power, e.

How should social and political o

of science and engineering

How do technologies transform the

Power and politics*

New technologies mean new cap

These capabilities are first ava

(cf., “The future is already here — it

How does this distribution of ca

How are data-empowered algorithms a

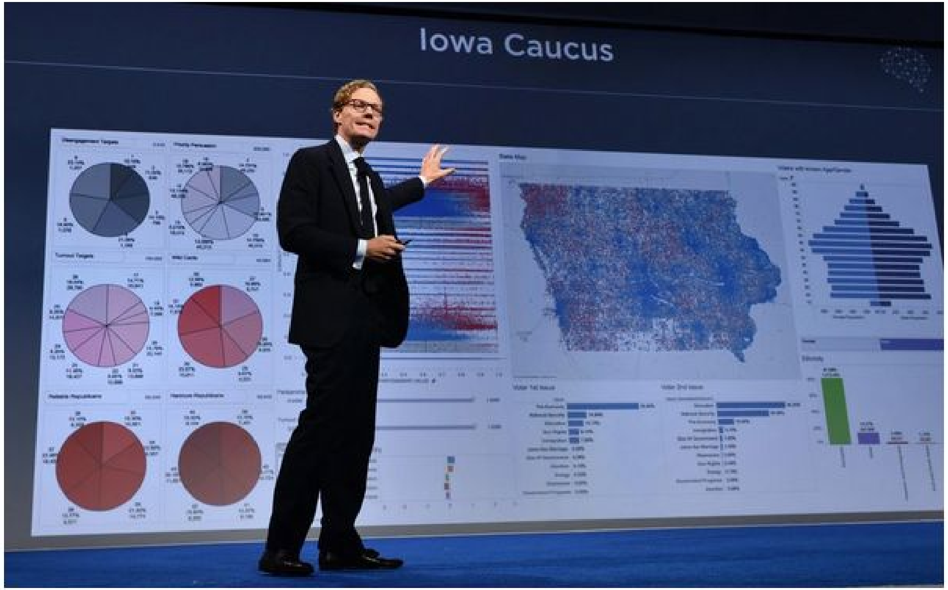
- of capability, and
- of reinforcing or distribut

* here meaning the dynamics of power, not to b

P o w e r   a n d   p o l i t i c s *

POLITICS | How Trump Consultants Exploited the Facebook Data o...

2076



Both Congress and the British Parliament have questioned Alexander Nix, chief executive of Cambridge Analytica, about the firm's activities. Bryan Bedder/Getty Images

*

This class: each week, re

Scientific and mathematical develo

Technologies and engineering

But also:

Driving forces: money, prestige, r

Truth, power, and data

This class: Every week

- Scientific and mathematical develo
- Technologies and engineering

But also:

- Driving forces: money, prestige, r
- Power, ethics, and data intensive

Weekly structure

Tuesday

Lecture and discussion

Expectation

arrive having done the week's readings

Thursday

Laboratory

Expectation

arrive with laptop re

Weekly structure

Tuesday

- Lessons from Slack
- “political” precondition
- intellectual advance
  - secondary readings
  - primary readings
- “political” postcondition
- what is contemporary analogue?
  - in capability
  - in politics
  - in data...

Thursday

Laboratory

...motivated by readi

...and presaging homew

Expectation

arrive with laptop re

“political” meaning

human implications and benefits

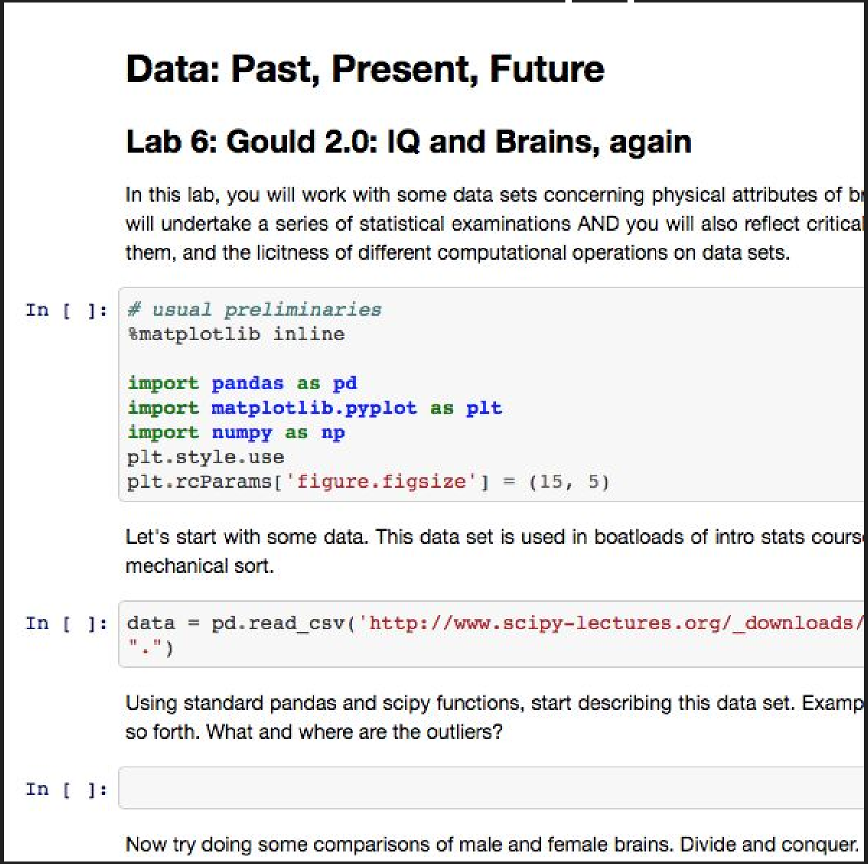
Jupyter notebooks: code without

USING JU

Using

platform

With everything you'll



Data: Past, Present, Future

Lab 6: Gould 2.0: IQ and Brains, again

In this lab, you will work with some data sets concerning physical attributes of br
will undertake a series of statistical examinations AND you will also reflect critical
them, and the licitness of different computational operations on data sets.

In [ ]:
```
# usual preliminaries
%matplotlib inline

import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
plt.style.use
plt.rcParams['figure.figsize'] = (15, 5)
```

Let's start with some data. This data set is used in boatloads of intro stats cours
mechanical sort.

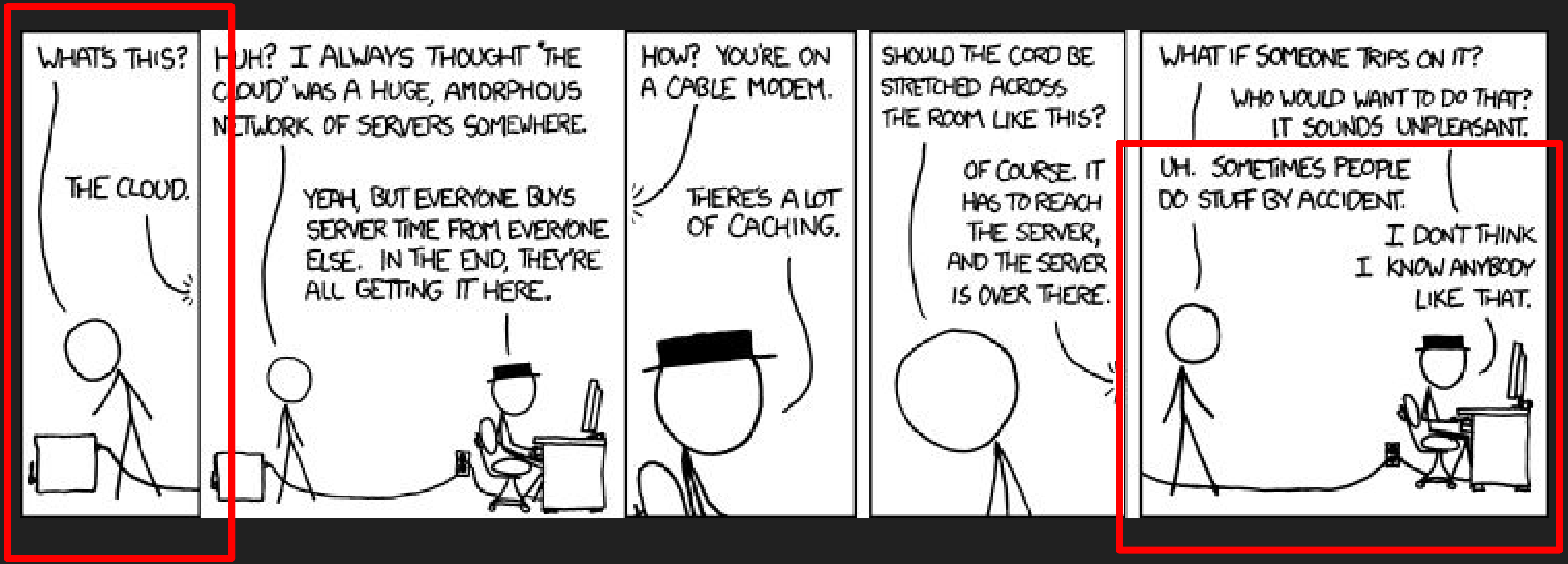
In [ ]:
```
data = pd.read_csv('http://www.scipy-lectures.org/_downloads/
".")
```

Using standard pandas and scipy functions, start describing this data set. Examp
so forth. What and where are the outliers?

In [ ]:

Now try doing some comparisons of male and female brains. Divide and conquer.

Colab: computing without computers



Colab: computing computers

https://colab.research.google.com/drive/1AB

Lab08b-20200402-perceptron.ipynb

File Edit View Insert Runtime Tools Help Last edited...

Later, the psychologist Frank Rosenblatt proposed a model to instantiate this in a mechanical computer -- not as "software" but as a specialized piece of electronics in the same way Claude Shannon's Theseus mouse was a special purpose computer.

Figure 4.8 Illustration of the Mark 1 perceptron hardware. The photograph on the left shows how the inputs were obtained using a simple camera system in which an input scene, in this case a printed character, was illuminated by powerful lights, and an image focussed onto a 20×20 array of cadmium sulphide photocells, giving a primitive 400 pixel image. The perceptron also had a patch board, shown in the middle photograph, which allowed different configurations of input features to be tried. Often these were wired up at random to demonstrate the ability of the perceptron to learn without the need for precise wiring, in contrast to a modern digital computer. The photograph on the right shows one of the racks of adaptive weights. Each weight was implemented using a rotary variable resistor, also called a potentiometer, driven by an electric motor thereby allowing the value of the weight to be adjusted automatically by the learning algorithm.

Check out his short paper https://psycnet.apa.org/fulltext/1959-09865-001.pdf or http://www.dtic.mil/docs/citations/AD0256582 for a detailed (>28MB) report.

Rosenblatt considered one output that took many signals as input.

A note on our 2 platforms (we'll re

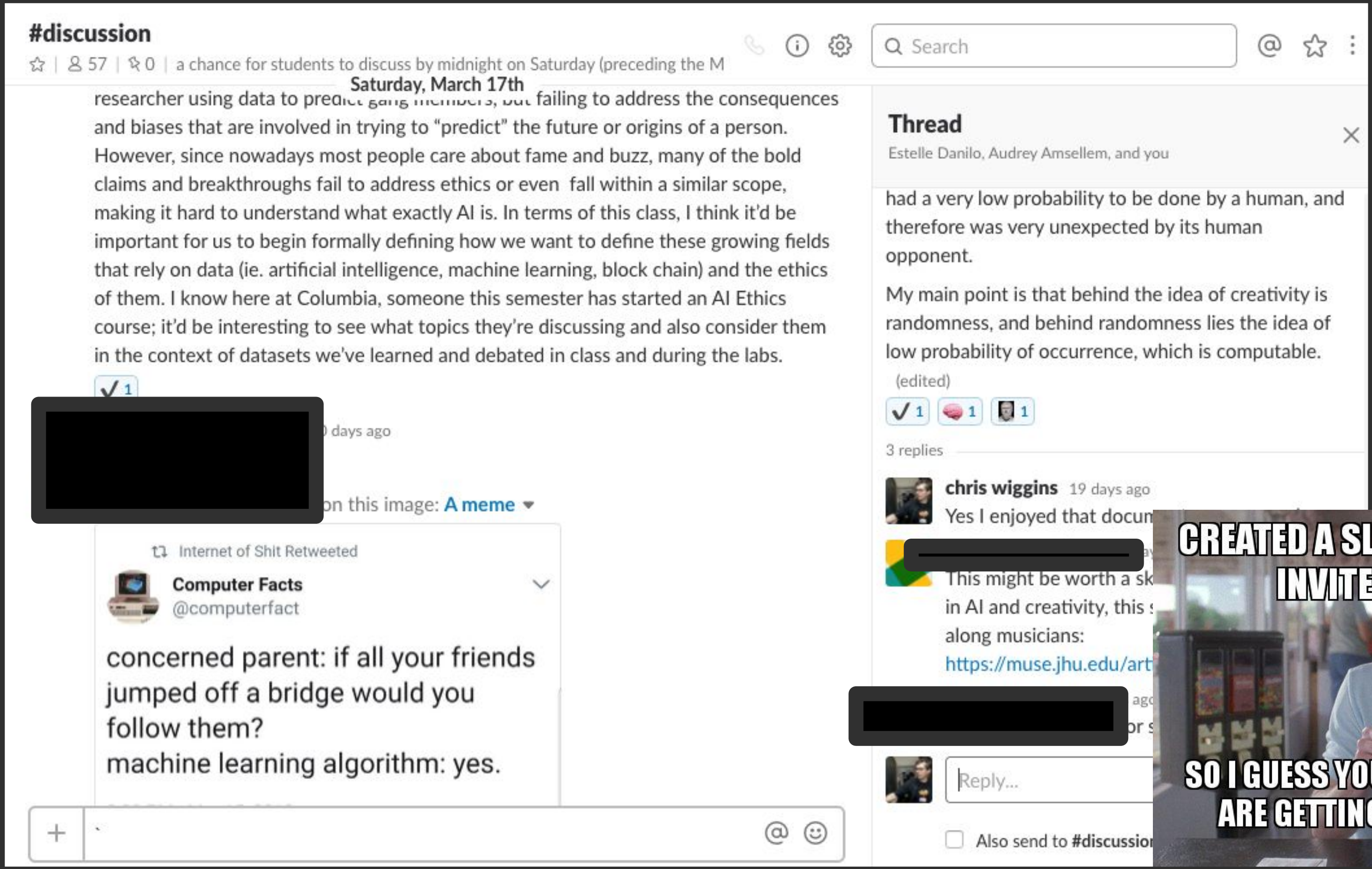
- You'll use Slack for #discussion and #labs
    (we'll use "#general" and #read
- You'll use Gradescope for your assignments g
  graders

Please ignore all else
- No Canvas,
- no Courseworks
- Github/course site is for people
  where canonical information is
    (and please let me know if it n

A note on grade distribution:

30% Homework (Gradescope)
30% Final Paper (Gradescope)
20% Op-Ed (Gradescope)
15% Discussion Participation (i
05% Lab Participation (in Slack

Slack-dom


#discussion
57 | 0 | a chance for students to discuss by midnight on Saturday (preceding the M
Search
Saturday, March 17th
researcher using data to predict gang members, but failing to address the consequences and biases that are involved in trying to "predict" the future or origins of a person. However, since nowadays most people care about fame and buzz, many of the bold claims and breakthroughs fail to address ethics or even fall within a similar scope, making it hard to understand what exactly AI is. In terms of this class, I think it'd be important for us to begin formally defining how we want to define these growing fields that rely on data (ie. artificial intelligence, machine learning, block chain) and the ethics of them. I know here at Columbia, someone this semester has started an AI Ethics course; it'd be interesting to see what topics they're discussing and also consider them in the context of datasets we've learned and debated in class and during the labs.
[redacted]
days ago
on this image: A meme
Internet of Shit Retweeted
Computer Facts
@computerfact
concerned parent: if all your friends jumped off a bridge would you follow them?
machine learning algorithm: yes.
Thread
Estelle Danilo, Audrey Amsellem, and you
had a very low probability to be done by a human, and therefore was very unexpected by its human opponent.
My main point is that behind the idea of creativity is randomness, and behind randomness lies the idea of low probability of occurrence, which is computable.
(edited)
3 replies
chris wiggins 19 days ago
Yes I enjoyed that docum
[redacted]
This might be worth a sk
in AI and creativity, this
along musicians:
https://muse.jhu.edu/art
[redacted]
Reply...
Also send to #discussion


CREATED A SLACK CHANNEL AND INVITED MEMBERS
SO I GUESS YOU COULD SAY THINGS ARE GETTING PRETTY SERIOUS
makeameme.org

Required work

Postings on readings in Slack each wee

Op-Ed (short but tight 750 words)

Participate in laboratory hours includ

Problem sets (extensions of lab work)

Final paper

Two tracks

more technical background track (60%)

- pursue a semester long project culminating in a 15 pp paper and any associated code
- complete 3 problem sets (the last 3 of the 5 problem sets)

more humanistic background

- write a 10 pp paper on a topic of choice
- complete 5 problem sets, th… sets will involve bot… and writing work

Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference

Two tracks / “Two Culture

“Two polar groups:
at one pole we have the lite
at the other scientists, …
Between the two a gulf of mu
incomprehension.”

- C.P. Snow, “the two culture

Two tracks

more humanistic background track (60%)

- write a 10pp paper on a topic of their choice
- complete 5 problem sets

more technical background

- write a 15pp paper on a topic of their choice
- complete 3 problem sets (t
  5 problem sets)

Two roads diverged in a wood, and I—
I took the one less traveled by,
And that has made all the difference

2 TAs (for writing); 3 gr

Graders:

@Fraye Beyene @William O'Brien @yamini

- All alumni of this class
- Fraye also a grader alumnus
- Python Primer: Sunday Jan 22 at 5 pm ! Loca

TAs:

@Jamie Nadel and James (slack name TBD)

- PhD students in History!

2 TAs (for writing); 3 gr

Graders:

@Fraye Beyene @William O'Brien @yamini

- All alumni of this class
- Fraye also a grader alumnus
- Python Primer: Sunday Jan 22 at 5 pm ! Loca

TAs:

@Jamie Nadel and James (slack name TBD)

- PhD students in History!

2 TAs (for writing); 3 gr

Graders:
@Fraye Beyene:
- If your *first* name starts with anything a
@William O’Brien:
- If your *first* name starts with anything a
@yamini
- If your *first* name starts with anything a

TAs:
@Jamie Nadel:
- If your *first* name starts with anything a
James:
- If your *first* name starts with anything a

No prerequisit

Registration

History-APMA UN2901 section 001

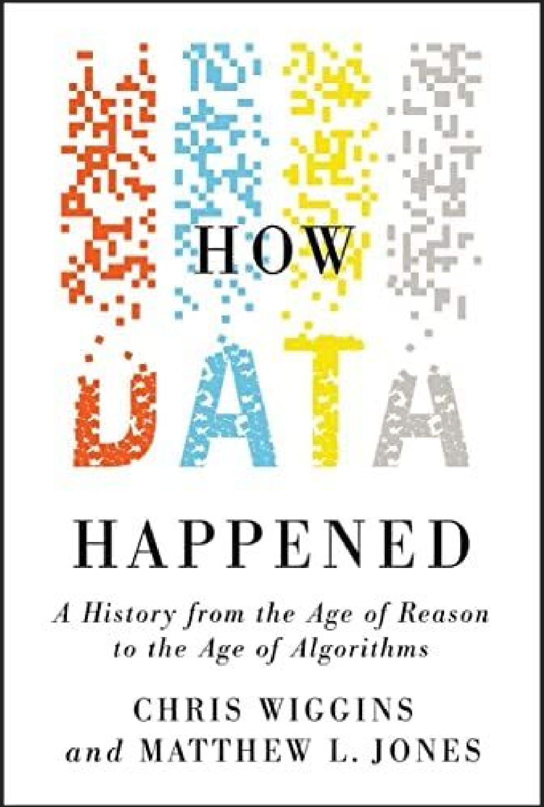
Call Number 12059

- The course satisfies one science req
  and counts towards "nontech elective
- This course satisfies Barnard's T&D
  Barnard thinking with historical con
- This course counts as a history cour
- NO PREREQUISITES AND NO PROGRAMMING
  REQUIRED!

Instructors' pe

required

Readings include new (Mar



First readings for next T

danah boyd & Kate Crawford

CRITICAL QUESTIONS FOR BIG DATA

Provocations for a cultural, technological, and scholarly phenomenon

2012

Hanna Wallach

CS professor. MSR NYC. Computational social science, data science, machine learning, parkour, Python, roller derby, sloths, vegan treats, visualization.

Dec 19, 2014 · 21 min read

2014

Big Data, Machine Learning, and the Social Sciences: Fairness, Accountability, and Transparency

This essay is a (near) transcript of a talk I recently gave at a NIPS 2014 workshop on “Fairness, Accountability, and Transparency in Machine Learning,” organized by Solon Barocas and Moritz Hardt.

Engineering the public: Big data, surveillance and computational politics

by Zeynep Tufekci

2014

First readings for next T

danah boyd & Kate Crawford

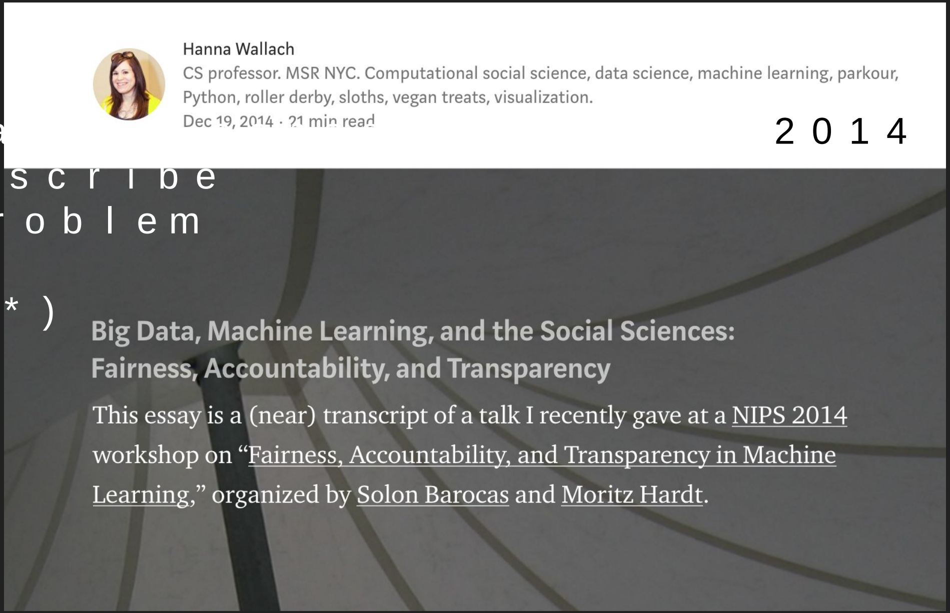
CRITICAL QUESTIONS FOR BIG DATA
Provocations for a cultural,
technological, and scholarly
phenomenon

2012

1. Big Data changes the de
2. Claims to objectivity a
3. Bigger data are not alwa
4. Taken out of context, B
5. Just because it is acce
6. Limited access to Big Da

First readings for next T

- Different ways of ma
  - e.g., predict, describe
- Not *just* a tech problem
- birth of “FATML”
  - (later ACM’s “FAT*)
  - (now ACM’s FAccT)



2014

2014

First readings for next T

- In view of Long history of politics +
- Provocation: what will happen when per
  architectures of advertising fuse with
  influence/information operations?

Engineering the public: Big data, surveillance and computational politics
by Zeynep Tufekci

2014

First readings for next T

danah boyd & Kate Crawford

CRITICAL QUESTIONS FOR BIG DATA
Provocations for a cultural, technological, and scholarly phenomenon

2012

Hanna Wallach
CS professor. MSR NYC. Computational social science, data science, machine learning, parkour, Python, roller derby, sloths, vegan treats, visualization.
Dec 19, 2014 · 21 min read

2014

Big Data, Machine Learning, and the Social Sciences: Fairness, Accountability, and Transparency

This essay is a (near) transcript of a talk I recently gave at a NIPS 2014 workshop on "Fairness, Accountability, and Transparency in Machine Learning," organized by Solon Barocas and Moritz Hardt.

Engineering the public: Big data, surveillance and computational politics
by Zeynep Tufekci

2014