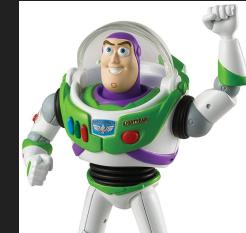


data: past, present, and future



chris wiggins + matt jones

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 X Power Editor - Manage Ads X

Secure | https://business.facebook.com/ads/manage/powereditor/manage/adsets

Search business Data: Past, Present, and Future 3 Help ?

Discard Changes Review Draft Items Settings

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 Spent
<input type="checkbox"/>	Ad Set - 1st PPF ad	Recently Completed	7 Link Clicks	258	399	\$0.57 Per Link Click	\$4.00 Lifetime	\$3.40
<input type="checkbox"/>	Ad Set - Ad B	Recently Completed	8 Link Clicks	259	381	\$0.50 Per Link Click	\$4.00 Lifetime	\$4.00
▶ Results from 2 ad sets ⓘ			15 Link Clicks	517 People	780 Total	\$0.53 Per Link Click		\$7.93 Total Spent

f

Pages X Power Editor - Manage Ads X

Secure | https://business.facebook.com/ads/manage/powereditor/manage/adsets/edit/

Power Editor Search business Data: Past, Present, and Future 1 Help ?

Creating Ad Set: test T-1 ad set

Detailed Targeting i INCLUDE people who match at least ONE of the following i

Demographics > Education > Schools
Columbia University

Interests > Additional Interests
Flat Earth

Add demographics, interests or behaviors | Suggestions | Browse

Audience Definition

Your audience selection is fairly broad.

Potential Reach: 810,000 people i

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.

No Results Found

Connections i

Placement

Saving to draft

2 items to review, including 1 new campaign Close



Ads Manager

Search business



Data: Past, Present, and Fu...

 Search 🔍

Traffic-20190121T14h54 > 2 Ad Sets > 2 Ads



Data: Past, Present, and Fu...

Active

In Split Test

...



Traffic-20190121T14h54

...

Ad Set for Ad B

...



Ad B

...

Ad Set for Ad A

...



Ad A

...

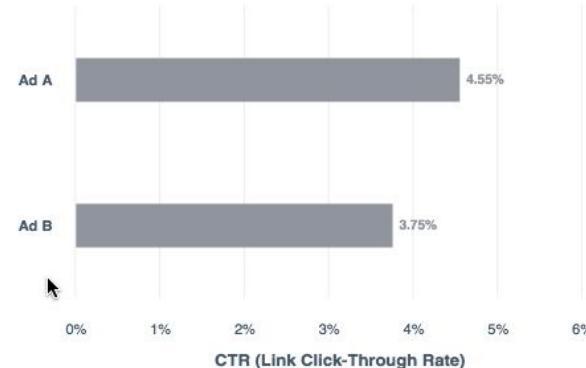
Split Test: Traffic-20190121T14h54 | Active

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.



Variable: Creative | Versions: 2 Ads | Total Budget: \$10.00, Even Split (50/50) | Objective: Traffic

Give Feedback

Ads Manager

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

Search business



Data: Past, Present, and Fu...



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** [Close](#)By clicking the "Publish" button, you agree to Facebook's [Terms and Advertising Guidelines](#).**SPONSORED**

Data: Past, Present, and ...

data: past, present, and future



*NO PREREQUISITES OR PROGRAMMING EXP...

[LEARN MORE >](#)

LIKE | COMMENT | SHARE

[Refresh preview](#) • [Report a problem with this preview](#)[Discard Draft](#)**Publish**

Ads Manager

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**

Search business



Data: Past, Present, and Fu...

**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** [Close](#)By clicking the "Publish" button, you agree to Facebook's [Terms and Advertising Guidelines](#).**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

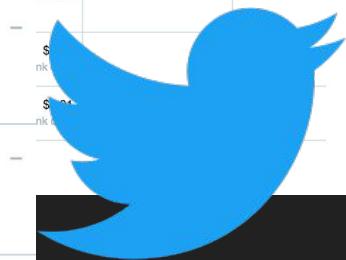
[Refresh preview](#) • [Report a problem with this preview](#)[Discard Draft](#)**Publish**

2018



Objective: All Status: All

Spend
\$25.00



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	-

Filters: Default

It's faster to monitor and optimize campaigns with our [Customization](#), [Automation](#), and [Export](#) capabilities.

Looking for? [Get help](#)

Campaigns Ad Groups Share

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

Spend \$25.00

Objective: All Status: All

Location ?

Location	Impressions	Spend	Results
Total for account	12,940	\$25.00	
Brazil	38	\$0.06	
Kuwait	18	\$0.00	

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.

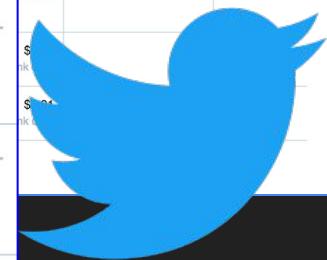
result	Daily budget	Re campaign
\$0.03 k click	\$5.00	
\$1.07 k click	-	
\$0.01 k click	-	
\$0.03 k click	\$5.00	
\$0.01 k click	-	

Kristi Hines @kikolani NEW: How to Create a Sales Page for Your Own Products [bit.ly/P8FM11](#) Guest post by @DominoConnect

Kristi Hines @kikolani The Ninja's Guide to Google Alerts [bit.ly/S4akiV](#) via @sejournal by @mattwoodwarduk

Kristi Hines @kikolani 9 Ways to Repurpose Content [bit.ly/QzP9lI](#) via @evergreensearch

Kristi Hines @kikolani 10 Sep How Google Panda & Penguin Affect Link Building (Past, Present & Future) [bit.ly/nVZG2M](#) via @cognitiveSEO





Search reports and help



HOME



CUSTOMIZATION



REAL-TIME



AUDIENCE



ACQUISITION



BEHAVIOR



CONVERSIONS



DISCOVER



ADMIN

Google Analytics Home



INTELLIGENCE

Users

476

↑833.3%

vs last 7 days

Sessions

505

↑770.7%

Bounce Rate

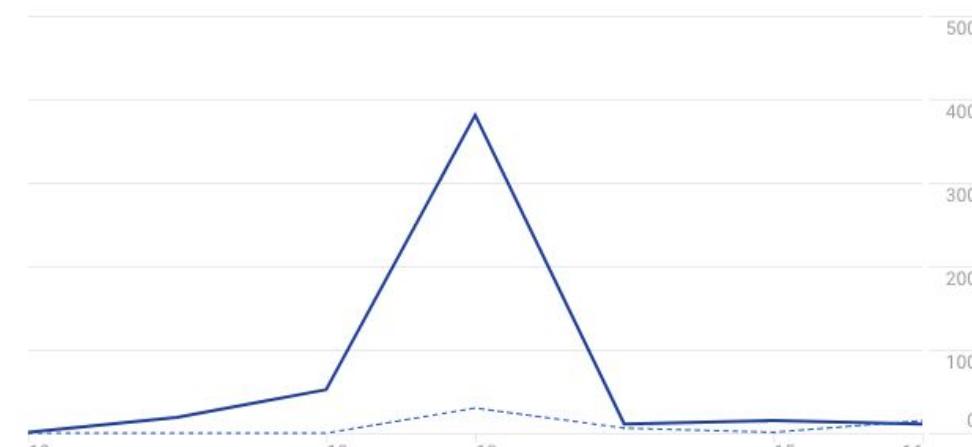
87.92%

↑6.2%

Session Duration

0m 13s

↓91.5%



Last 7 days ▾

AUG

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			



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			



Country / Medium	Acquisition			Behavior			Conversions			
	Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration	Goal Conversion Rate	Goal Completions	Goal Value	Goal % of Total
	5633 % of Total: 100.00% (563)	93.43% Avg. New Session: 0.00%	526 % of Total: 100.00% (526)	87.39% Avg. Bounce Rate: 0.00%	1.24 Avg. Pages/Session: 0.00%	00:00:29 Avg. Session Duration: 0.00%	0.00% Avg. Goal Conversion Rate: 0.00%	0 % of Total: 0.00% (0)	\$0.00 Avg. Goal Value: \$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
1. indonesiæ / direct	271 (48.19%)	how did this end up in my news feed? why?								
2. t.co / referral United States	149 (26.47%)	0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00%								
3. ads-bidder-api.twitter.com / referral Malaysia	81 (14.39%)	0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00%								
4. m.facebook.com / referral Thailand	8 (1.60%)	0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00%								
5. l.facebook.com / referral Australia	7 (1.24%)	0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00%								
6. datascience.columbia.edu / referral United Kingdom	7 (1.24%)	0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00%								
7. google / organic India	6 (1.07%)	0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00% 0.00%								
8. facebook.com / referral Switzerland	4 (0.71%)	100.00% (0.71%)	(0.76%)	100.00%	1.00	00:00:00	0.00% 0.00%	0 (0.00%)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
9. ad-review-tool.twitter.biz / referral Peru	4 (0.71%)	100.00% (0.71%)	(0.76%)	75.00%	1.25	00:00:17	0.00% 0.00%	0 (0.00%)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
10. adwords.google.com / referral Turkey	4 (0.71%)	100.00% (0.71%)	(0.76%)	50.00%	1.95	00:00:06	0.00% 0.00%	0 (0.00%)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)

Rows 1 - 10 of 12

Country / Medium	Acquisition			Behavior			Conversions			
	Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration	Goal Conversion Rate	Goal Completions	Goal Value	Goal % of Total
	5633 % of Total: 100.00% (563)	93.43% Avg. New Session: 0.00%	526 % of Total: 100.00% (526)	87.39% Avg. Bounce Rate: 0.00%	1.24 Avg. Pages per View: 0.00	00:00:29 Avg. Session Duration: 0.00	0.00% Avg. Goal Conversion Rate: 0.00%	0 % of Total: 0.00% (0)	\$0.00 Avg. Goal Value: \$0.00	\$0.00 % of Total: 0.00% (0)
1. (direct) Indonesia	271 (48.19%)									
2. t.co / referral United States	149 (26.47%)									
3. ads-bidder-api.twitter.com / referral Malaysia	81 (14.39%)									
4. m.facebook.com / referral Thailand	8 (1.46%)									
5. l.facebook.com / referral Australia	7 (1.24%)									
6. datascience.columbia.edu / referral United Kingdom	7 (1.24%)									
7. google / organic Indie	6 (1.07%)									
8. facebook.com / referral Switzerland	4 (0.71%)	100.00% (0.71%)	(0.76%)	100.00% (0.71%)	1.00	00:00:00	0.00% (0.00%)	0 (0.00%)	\$0.00 (0.00%)	\$0.00 (0.00%)
9. ad-review-tool.twitter.biz / referral Peru	4 (0.71%)	100.00% (0.71%)	(0.76%)	75.00% (0.71%)	1.25	00:00:17	0.00% (0.00%)	0 (0.00%)	\$0.00 (0.00%)	\$0.00 (0.00%)
10. adwords.google.com / referral Turkey	4 (0.71%)	100.00% (0.71%)	(0.76%)	50.00% (0.71%)	1.95	00:00:06	0.00% (0.00%)	0 (0.00%)	\$0.00 (0.00%)	\$0.00 (0.00%)

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?

Country / Medium	Acquisition			Behavior			Conversions			
	Sessions	% New Sessions	New Users	Bounce Rate	Pages / Session	Avg. Session Duration	Goal Conversion Rate	Goal Completions	Goal Value	Goal % of Total
	5633 % of Total: 100.00% (563)	93.43% Avg. New Session: 0.00%	526 % of Total: 100.00% (526)	87.39% Avg. Bounce Rate: 0.00%	1.24 Avg. Pages per View: 0.00	00:00:29 Avg. Session Duration: 0.00	0.00% Avg. Goal Conver. 0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 Avg. Goal Value: \$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
And...										
1. indonesiæ / referral	271 (48.19%)						0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
2. t.co / referral United States	149 (26.47%)						0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
3. ads-bidder-api.twitter.com / referral Malaysia	81 (14.39%)						0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
4. m.facebook.com / referral Thailand	8 (1.46%)						0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
5. l.facebook.com / referral Australia	7 (1.24%)						0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
6. datascience.columbia.edu / referral United Kingdom	7 (1.24%)						0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
7. google / organic India	6 (1.07%)						0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
8. facebook.com / referral Switzerland	4 (0.71%)	100.00% (0.71%)	(0.76%)	100.00%	1.00	00:00:00	0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
9. ad-review-tool.twitter.biz / referral Peru	4 (0.71%)	100.00% (0.71%)	(0.76%)	75.00%	1.25	00:00:17	0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)
10. adwords.google.com / referral Turkey	4 (0.71%)	100.00% (0.71%)	(0.76%)	50.00%	1.95	00:00:06	0.00% (0.00%)	0 % of Total: 0.00% (0)	\$0.00 (\$0.00)	\$0.00 % of Total: 0.00% (0)

Rows 1 - 10 of 12

**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

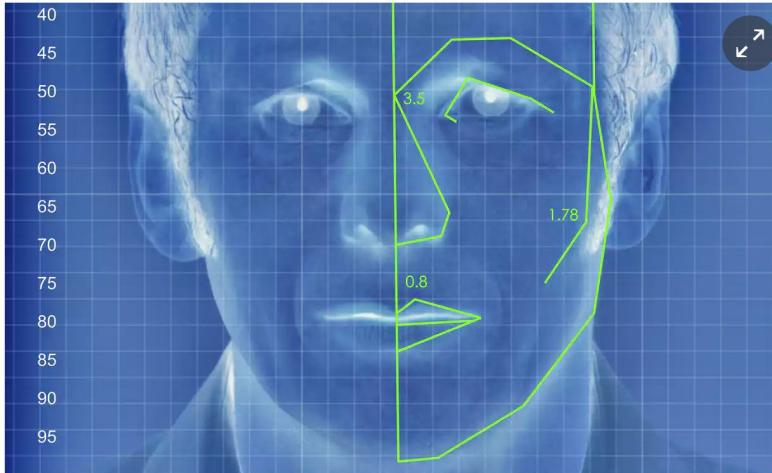


Illustration: Alamy

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.

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

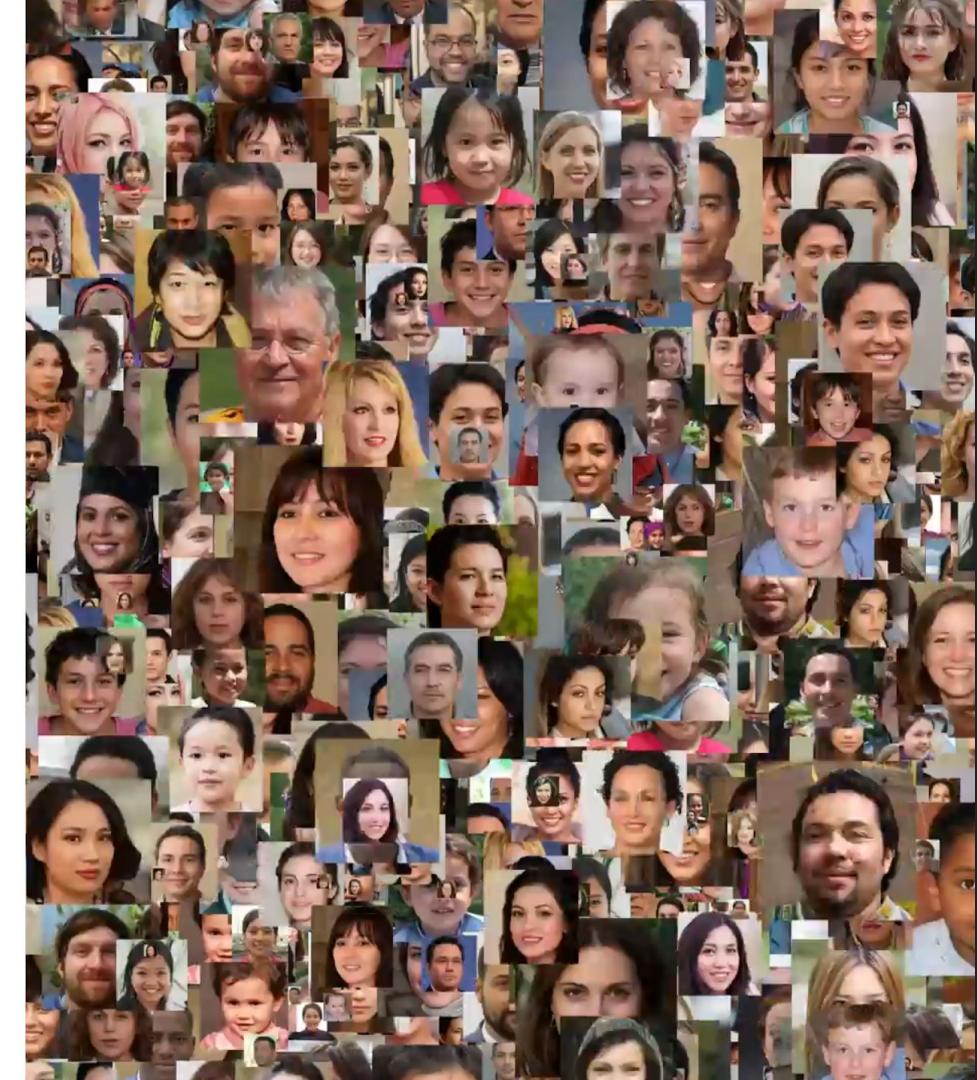
Category: Project

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 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 *Times*, 2020-01-18



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.

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 and all periods of recorded human history, people share the belief that the face alone suffices to reveal innate traits of a 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

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 ▾

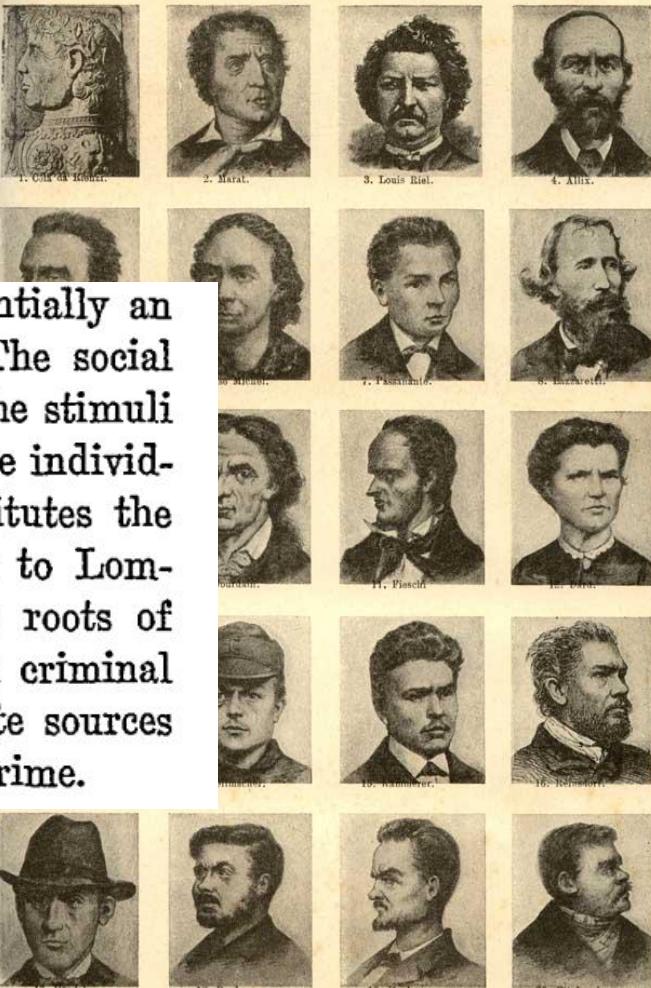


the onion dot com

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

We've been here before

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 Aguera y Arcas, [Alexander Todorov](#) and [Margaret Mitchell](#)

We've been here before



**Do
just**

by Bla

Medium

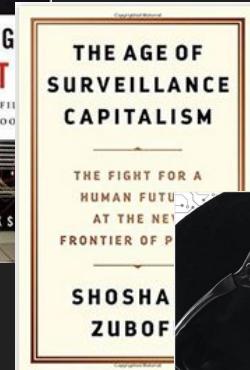
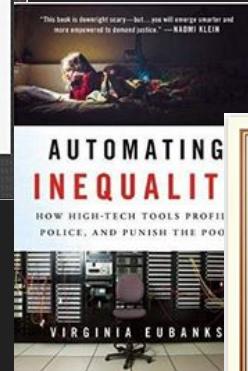
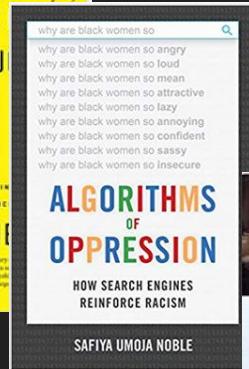
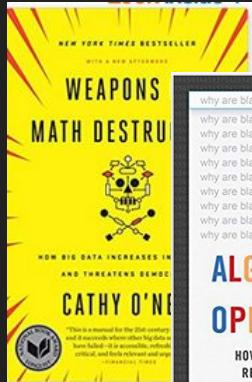
[Sign in](#)

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.

ne of the

or

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



2017-09-05: cathy o'neil

2018-01-08: safiya noble

2018-01-23: virginia eubanks

2019-01-15: shoshana zuboff

2019-06-17: ruha benjamin

something is wrong on the
internet

Statistical sciences always political

Dream of sciences of social difference

Central to development of

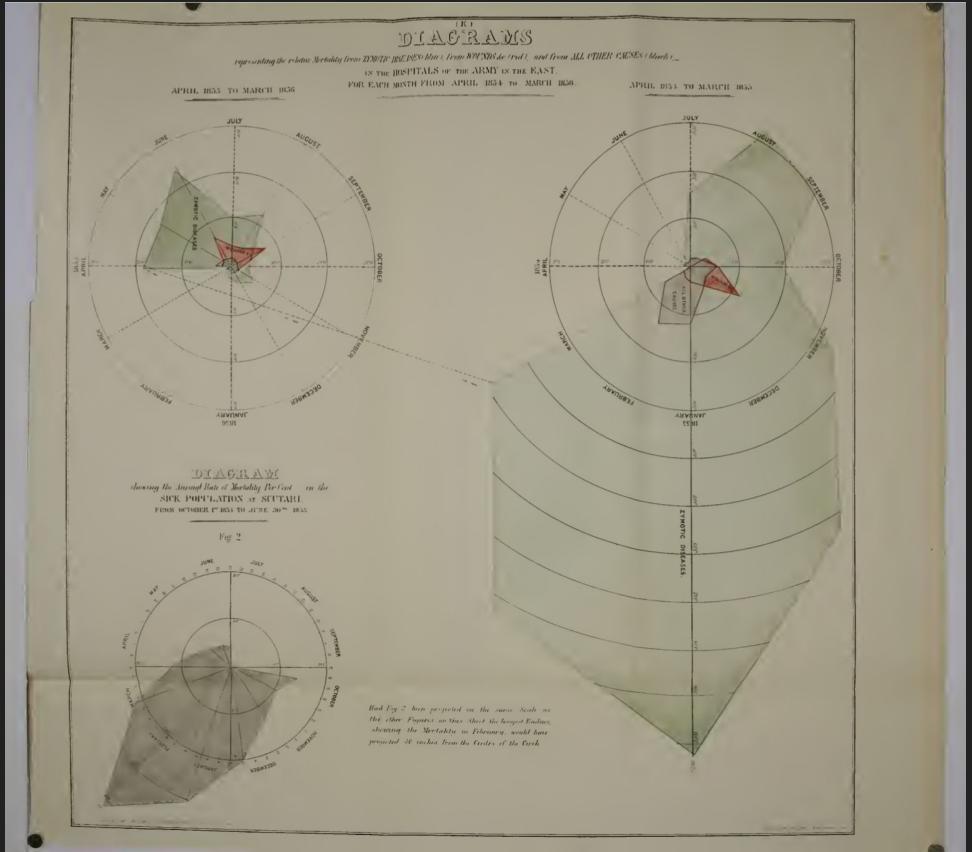
Statistics

And the

Data sciences

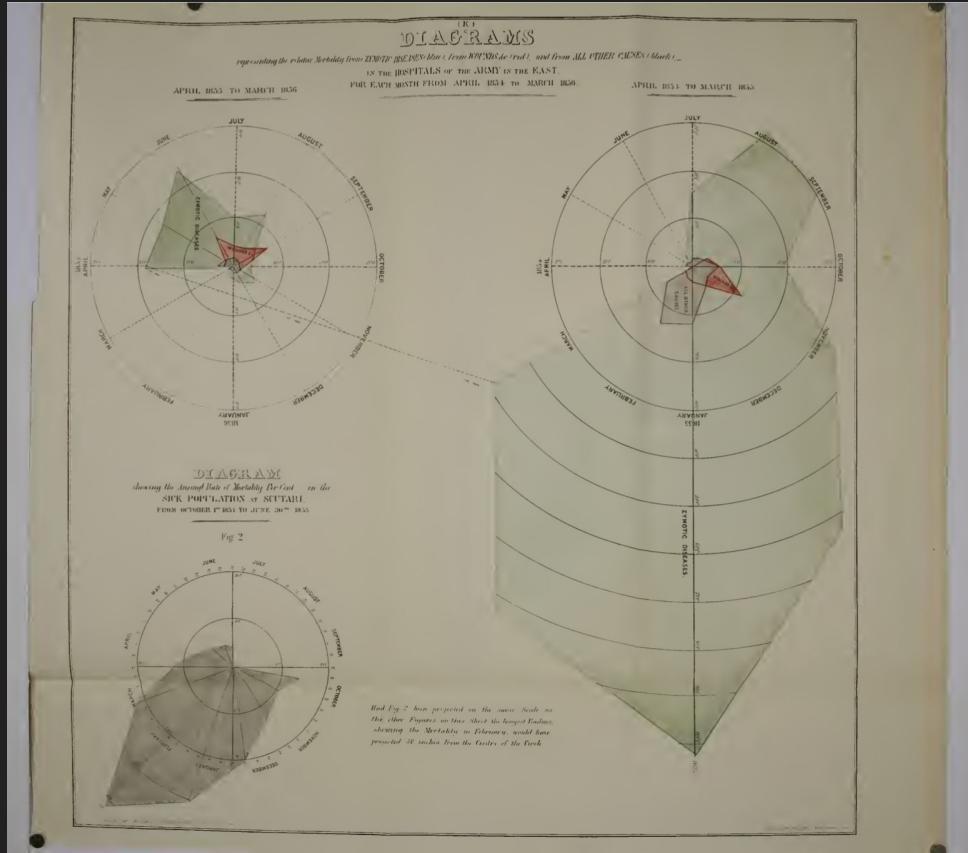
Florence Nightingale & Data Visualization

“Experience has shown that without special information and skilful application of the resources of science in preserving health, the drain on our home population must exhaust our means. The introduction, therefore, of a proper sanitary system into the British army is of essential importance to the public interests.”



Florence Nightingale & Data Visualization

“Upon the British race alone the integrity of that empire at this moment appears to depend. The conquering race must retain possession.”



Every week

Scientific and mathematical development

Technologies and engineering

Driving forces: money, prestige, resources, Imperial competition

Power, ethics, and data intensive knowledge

Tech story: three chronological stages

Data and Math

Data and Engineering

Data and Technology

Data technologies

Census and government survey

Information processing machines and digital computers

Always on network infrastructure



Power

How should social and political order be organized on basis
of science and engineering?

How do technologies transform the social and political order?

Phillip Rogaway, "The Moral Character of Cryptographic 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

Power and politics*

New technologies mean new capabilities

These capabilities are first available to those in power

(cf., “The future is already here — it's just not very evenly distributed.” --Gibson)

How does this distribution of capability reorder power?

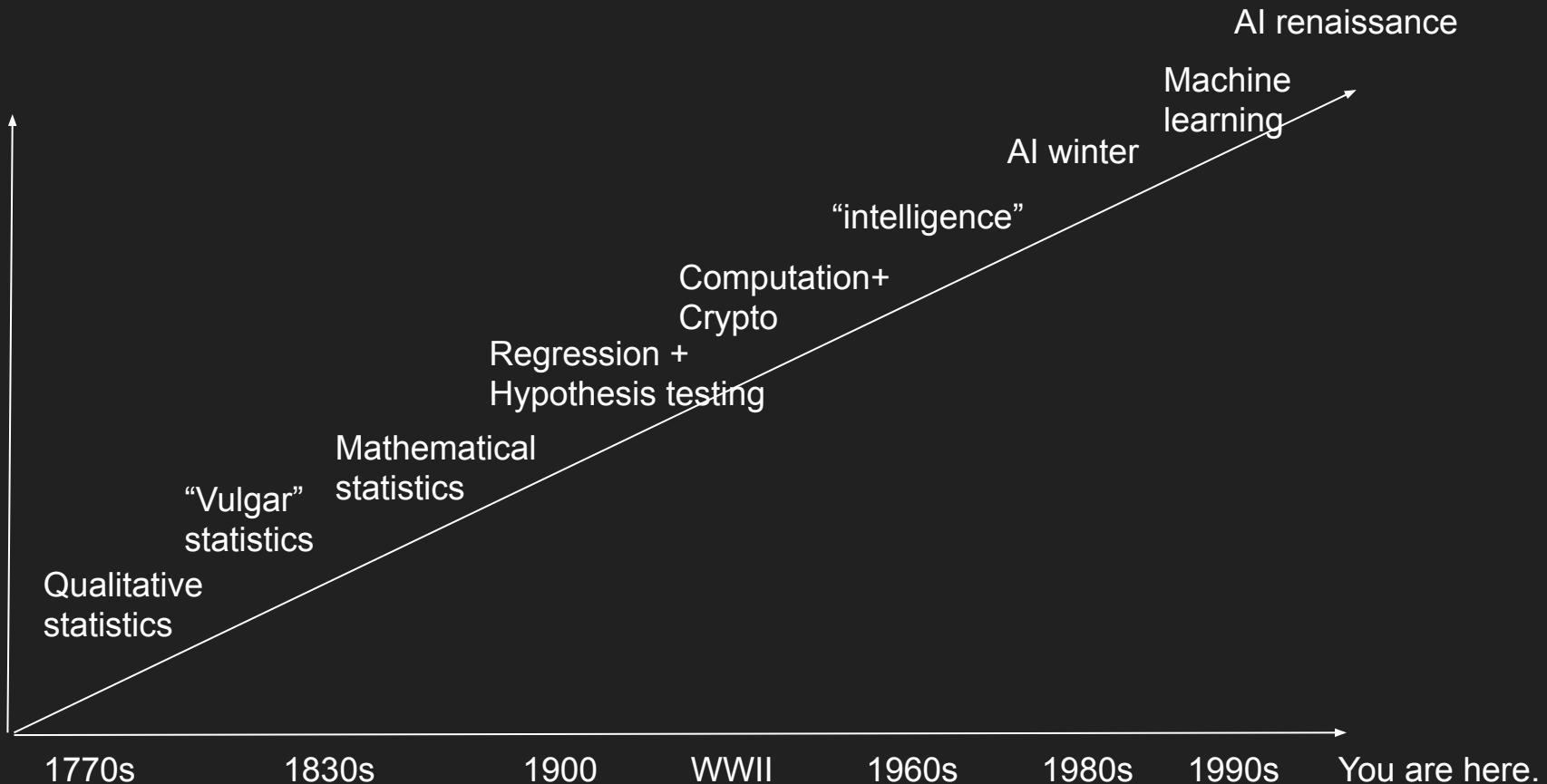
How are data-empowered algorithms an example of this dynamic

- of capability, and
- of reinforcing or distributing power?

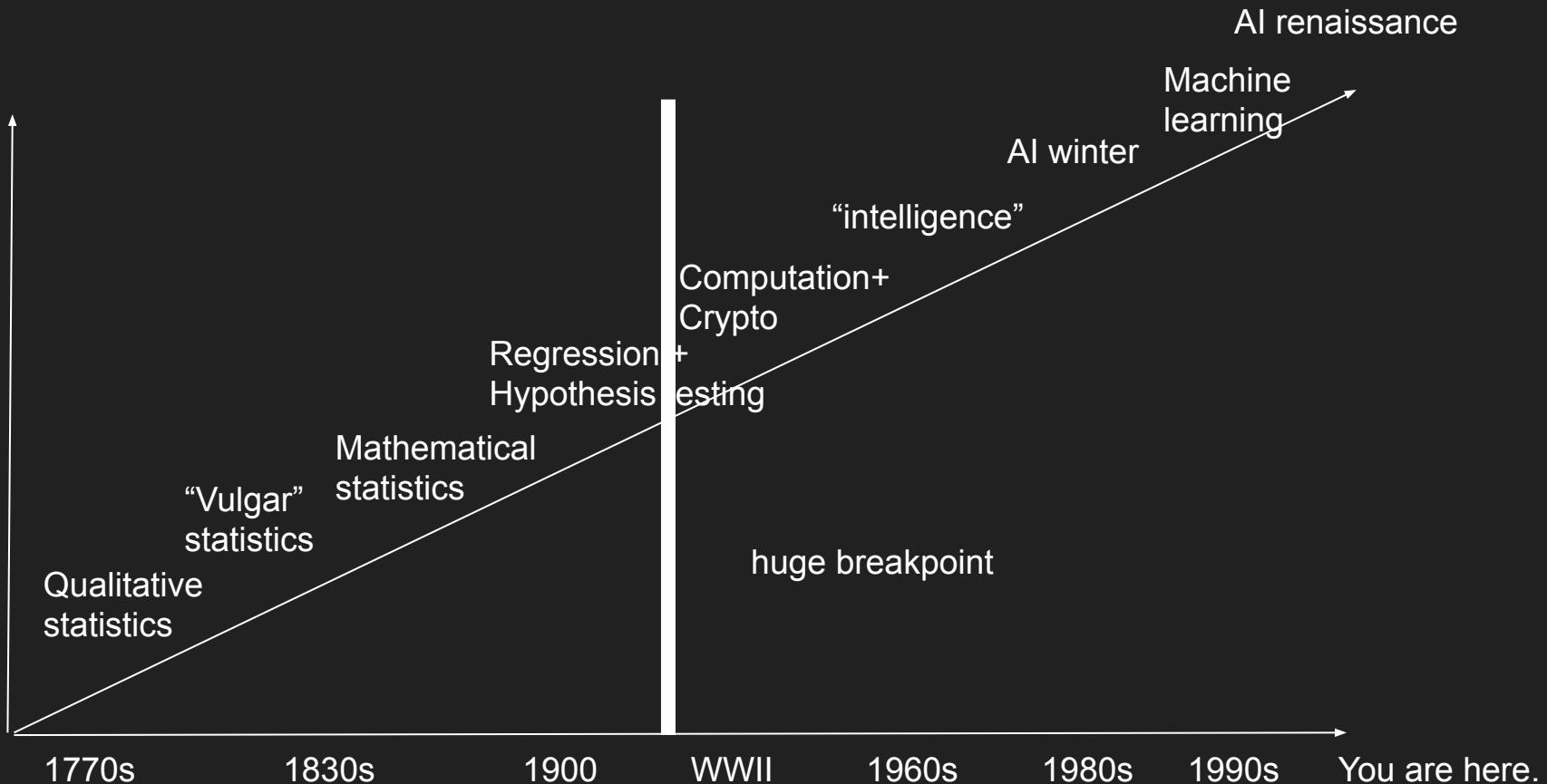
* *politics* here meaning the dynamics of power, not to be confused with “voting”

Arc of class

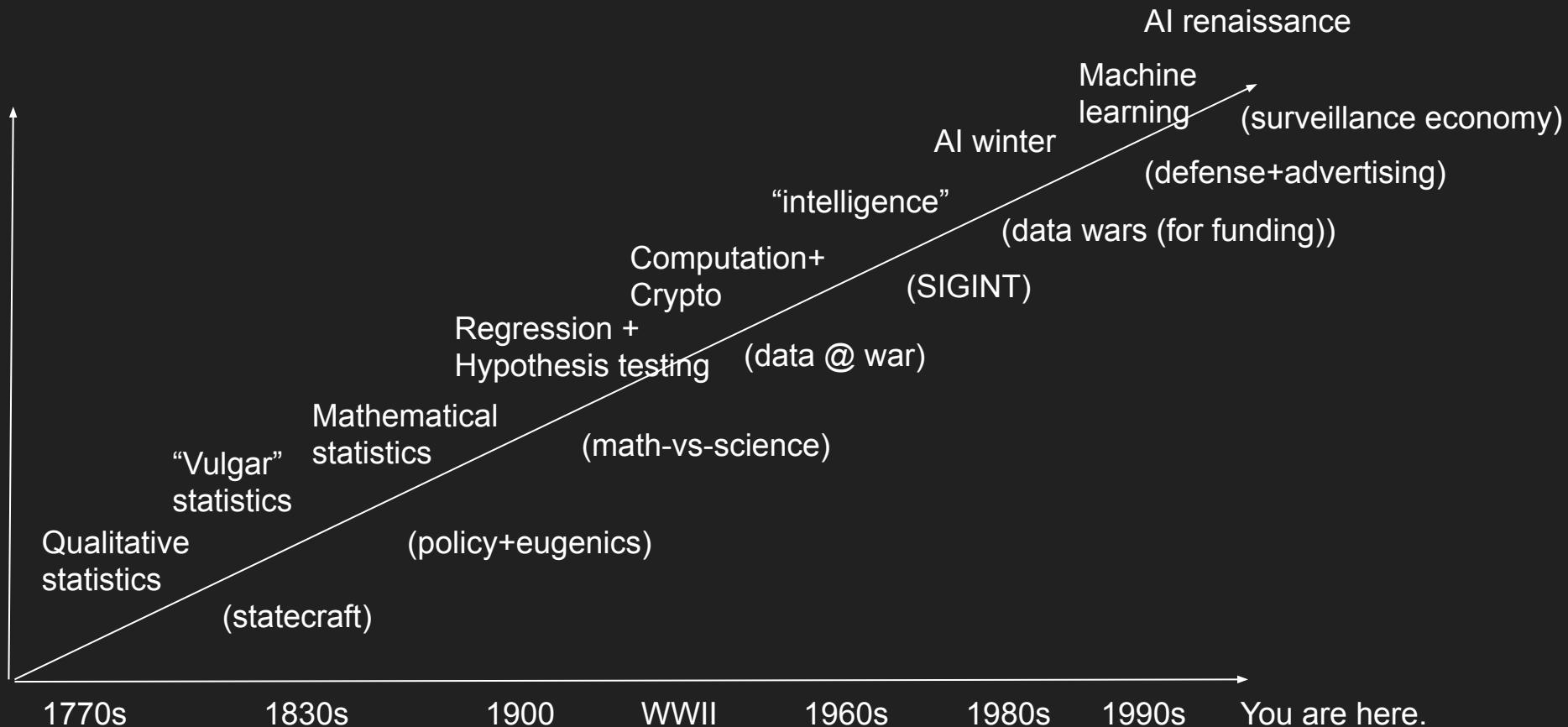
data 1770s-present: capabilities



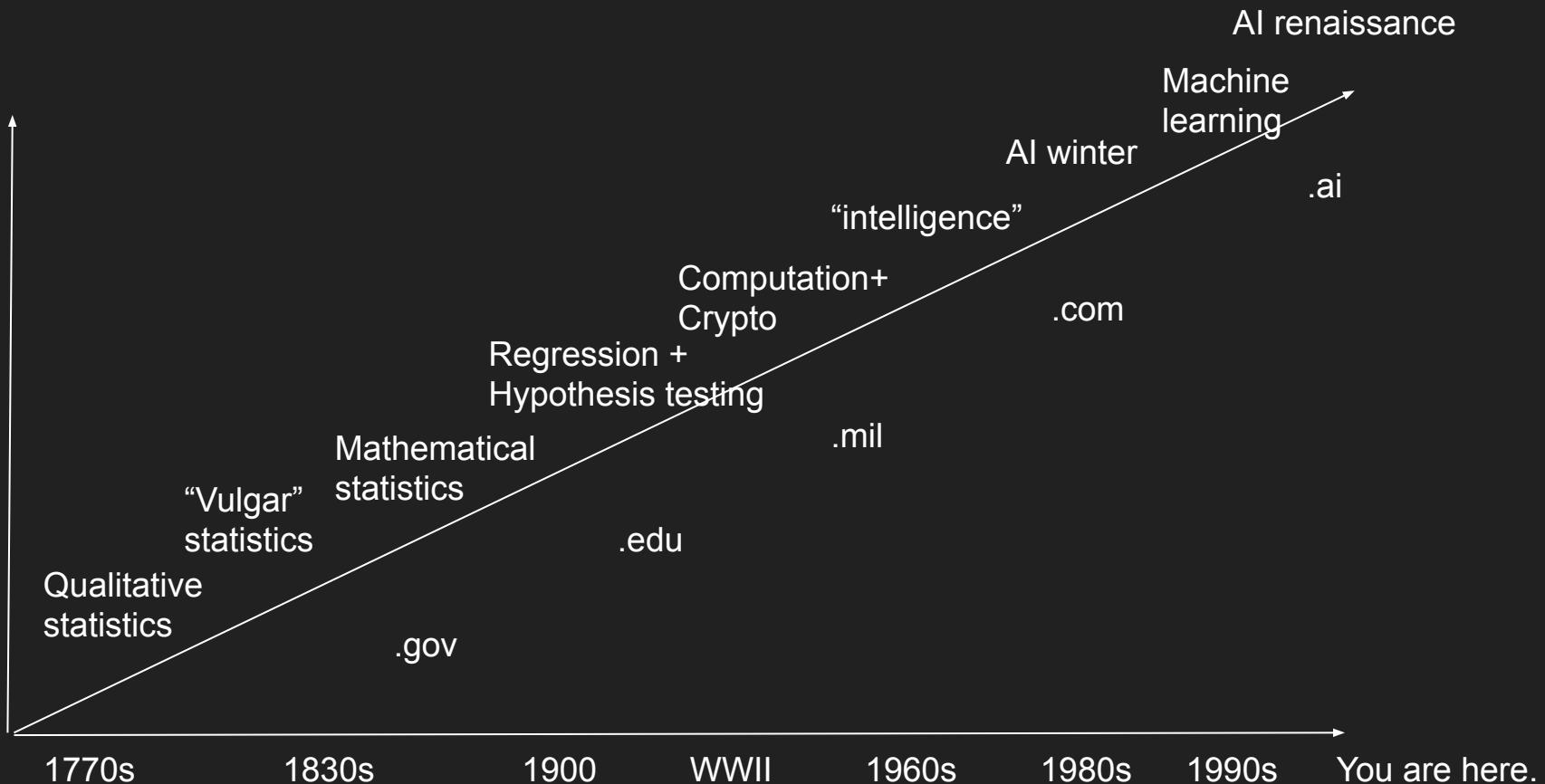
data 1770s-present: capabilities



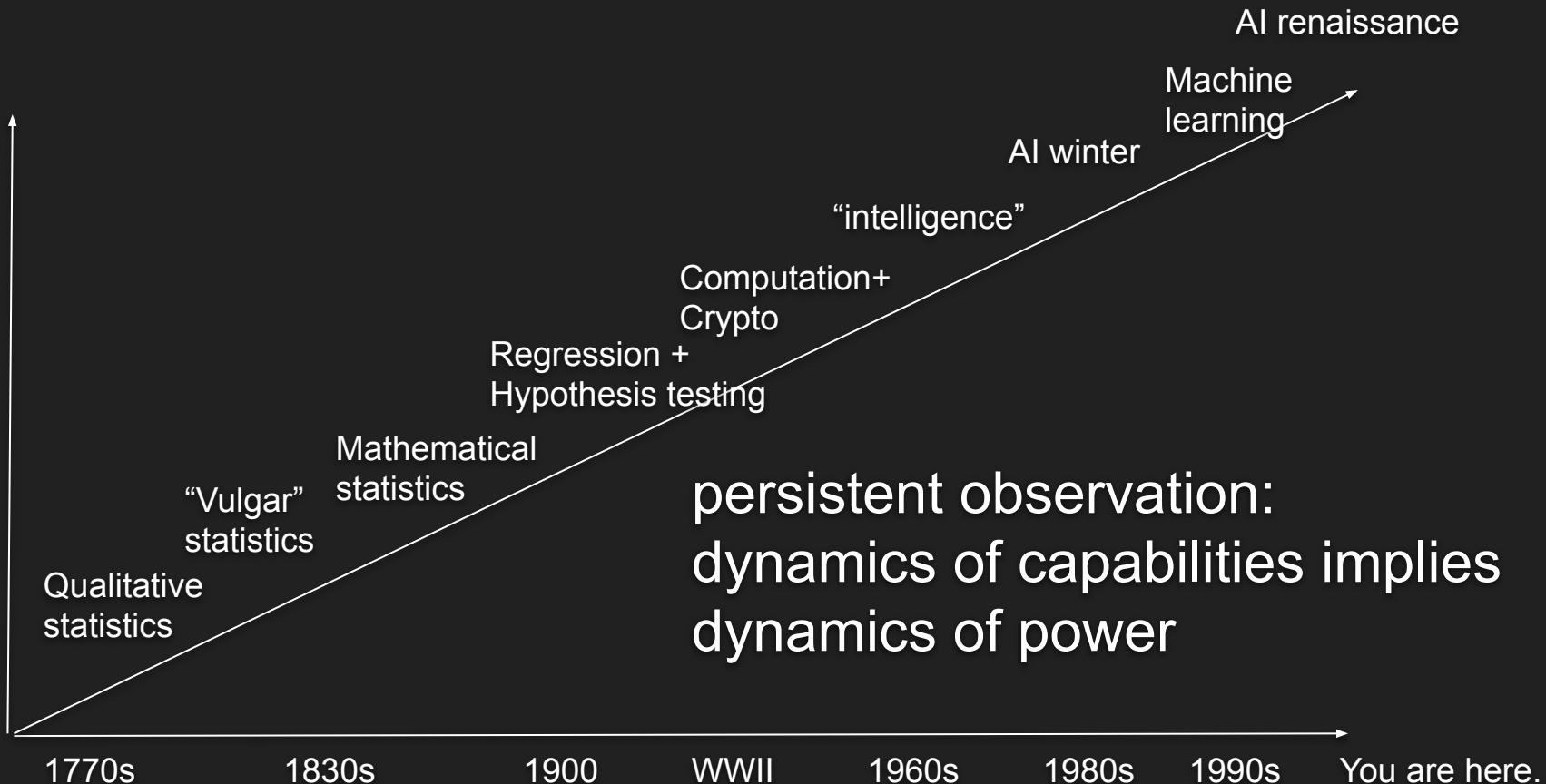
data 1770s-present: capabilities & intents



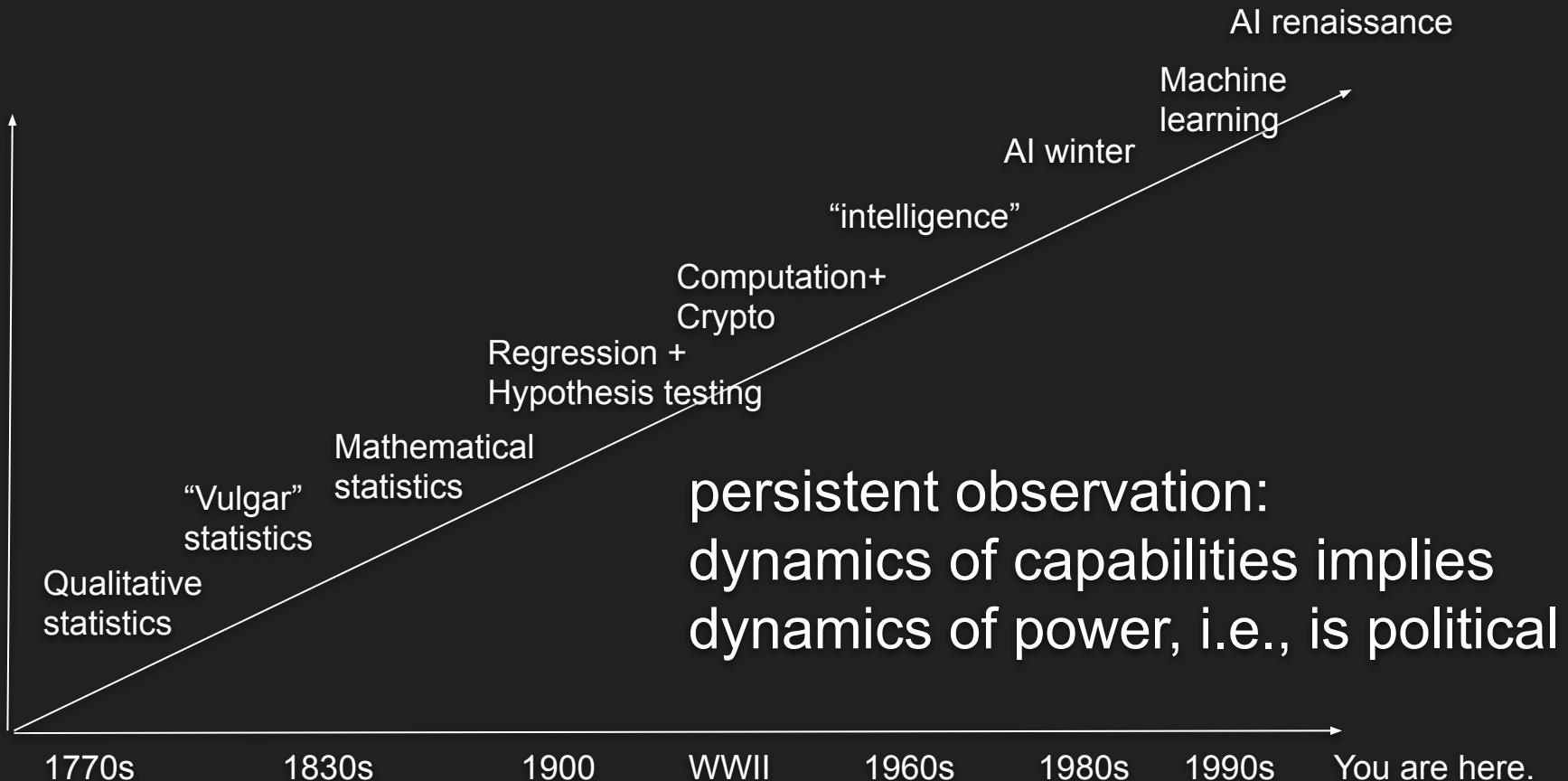
data 1770s-present: capabilities & intents



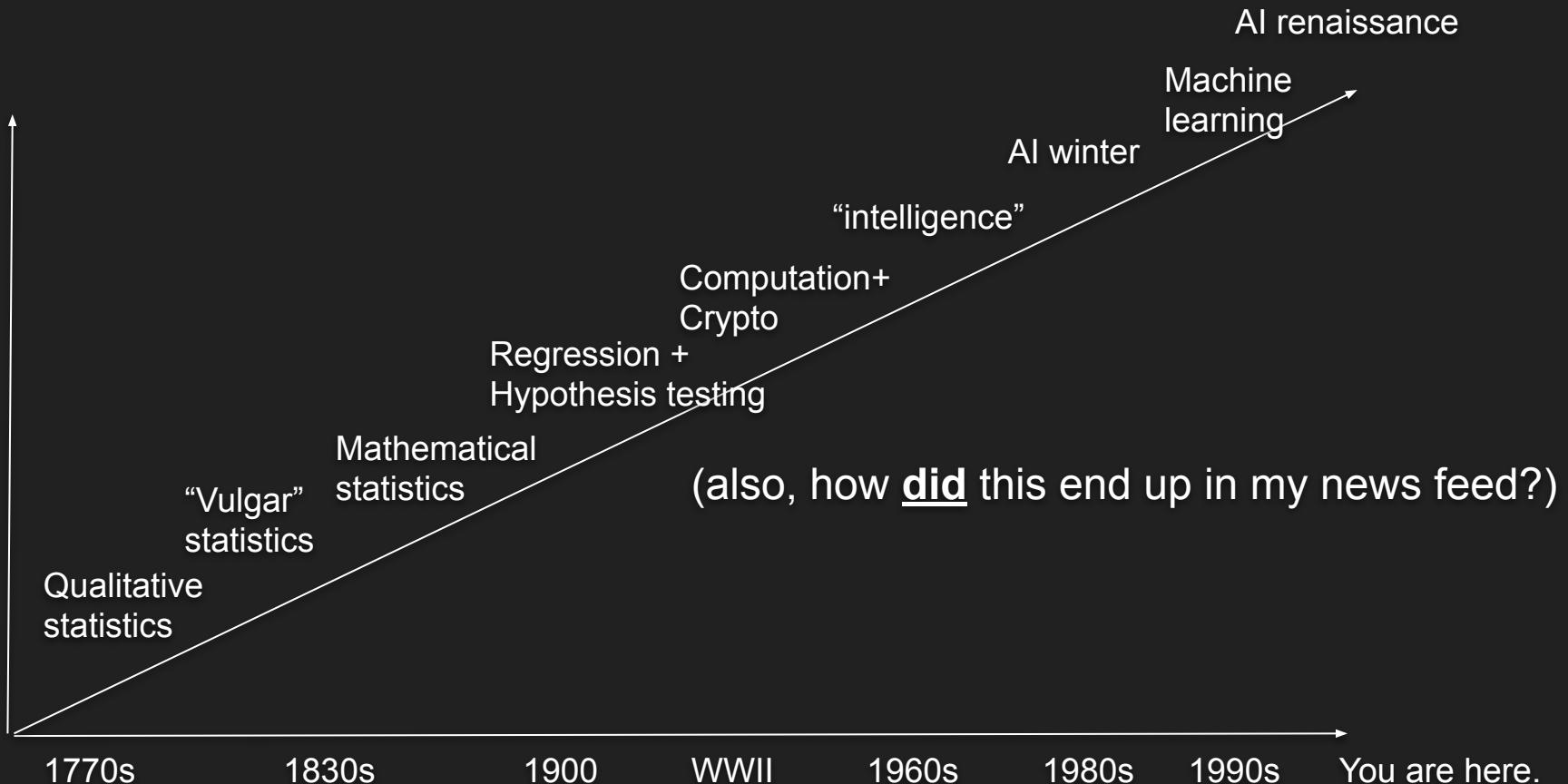
data 1770s-present: capabilities & intents



data 1770s-present: capabilities & intents



data 1770s-present: capabilities & intents



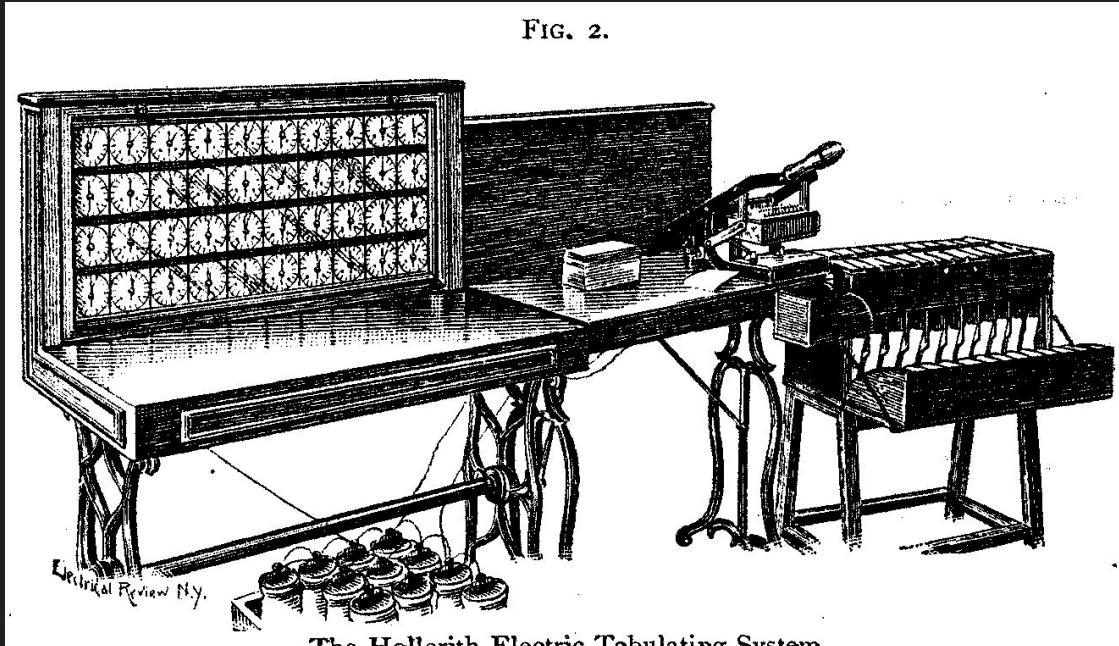
How we know what the state of the state is

Tabelle Schem. B.
Verzeichnis der Gestorbenen nach dem Alter und Jahren.
Zusammen im Jahr 1887.

Alter und Jahre.	Männl. den Geschl.	Weibl. den Geschl.	Summa von beiden.
Erwachsene ...	266	204	470
bis zum ersten Jahr, vom ersten bis zum zehn. incl.	126	162	288
6 - 10 - - -	742	724	1466
11 - 15 - - -	218	203	421
16 - 20 - - -	114	99	213
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 - - -	193	220	413
56 - 60 - - -	258	301	559
61 - 65 - - -	321	396	697
66 - 70 - - -	298	354	652
71 - 75 - - -	293	273	566
76 - 80 - - -	223	266	429
81 - 85 - - -	110	126	236
86 - 90 - - -	76	92	168
91 - 95 - - -	72	19	91
96 - 99 - - -	4	1	5
100 Jahr - - - -	1	-	-
101 - - - -	-	-	-
102 - &c. - - - -	-	-	-
Summa	5438	5918	10356

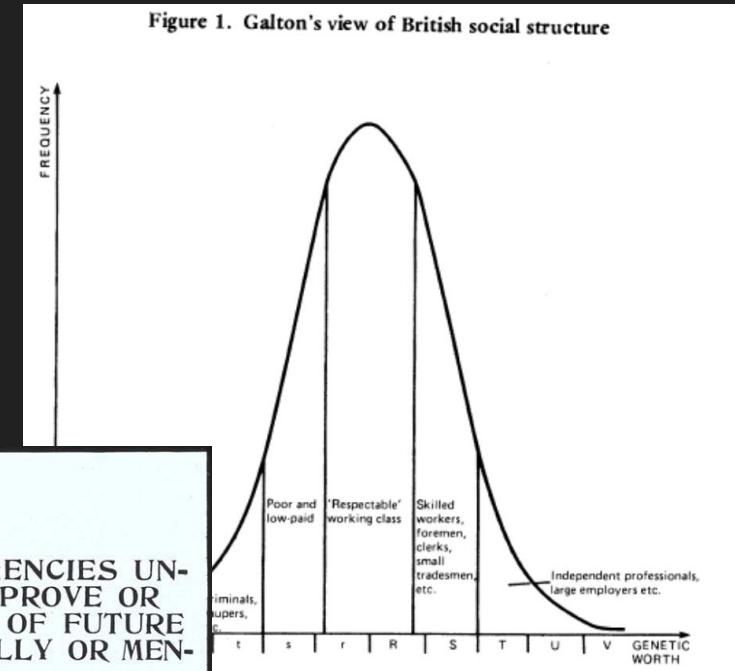
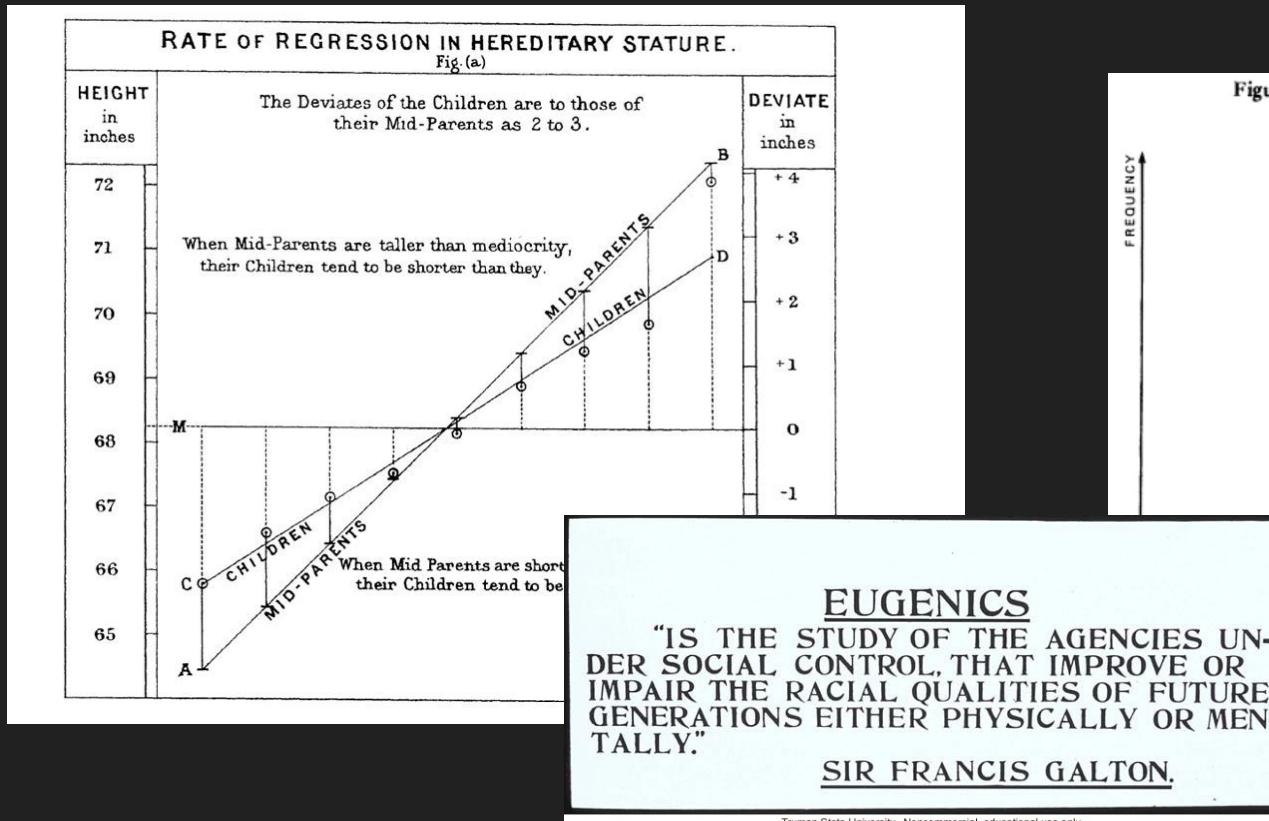
Zusammenfassung 1. Wenn keine Ziffer ist eine Frau von ganz gesetztem Grade und die Zahl steht auf Kreuzchen ist sie ein Jahr als erwachsen. Zusammenfassung 2. Ein Zeichenmann ist bei jeder als erwachsen, während kein völliges Geschlecht einer jungen Dame ist am Ende erhalten. Dergleiche Zusammenfassungen geben die Zahl nicht an die Hand.

FIG. 2.



The Hollerith Electric Tabulating System.

Regression, correlation and eugenics



e.g., week 4 regression & quantitative racism



Stephen Jay Gould
The Mismeasure of Man
A brilliant and controversial study
of intelligence testing
'Superlative'—*Nature*

Measuring Heads

Paul Broca and the Heyday of Craniology

No rational man, cognisant of the equal, still less the superior, will be true, it is simply incredible that our prognathous relative is an oppressor, he will be able to conquer a smaller-jawed rival, in a conflict of 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 & quantitative racism

Data: Past, Present, Future | Lab 4 | 2/14/2019

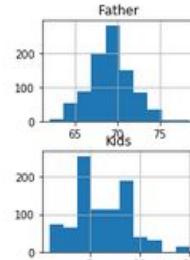
describing and predicting: Galton, regression, inventing error, sum

Galton and regression

Galton's analysis "gives the number and precision [see Plate IX, fig. 6]

```
In [7]: heights.hist()
```

```
Out[7]: array([[[<matplotlib.axes._subplots.AxesSubplot object at 0x0000000000000000>, <matplotlib.axes._subplots.AxesSubplot object at 0x0000000000000000>, <matplotlib.axes._subplots.AxesSubplot object at 0x0000000000000000>, <matplotlib.axes._subplots.AxesSubplot object at 0x0000000000000000>, <matplotlib.axes._subplots.AxesSubplot object at 0x0000000000000000>], dtype=object)
```



```
In [28]: # plot fit line  
plt.scatter(x, y, color='black')  
plt.plot(x, skl_lm.predict(x), color='blue', linewidth=1)  
plt.show()
```



Now it's your turn!

can you the regression for

- 1. everybody and his/her mother?
- 2. males and fathers

Experimental design, hypothesis test, decision theory

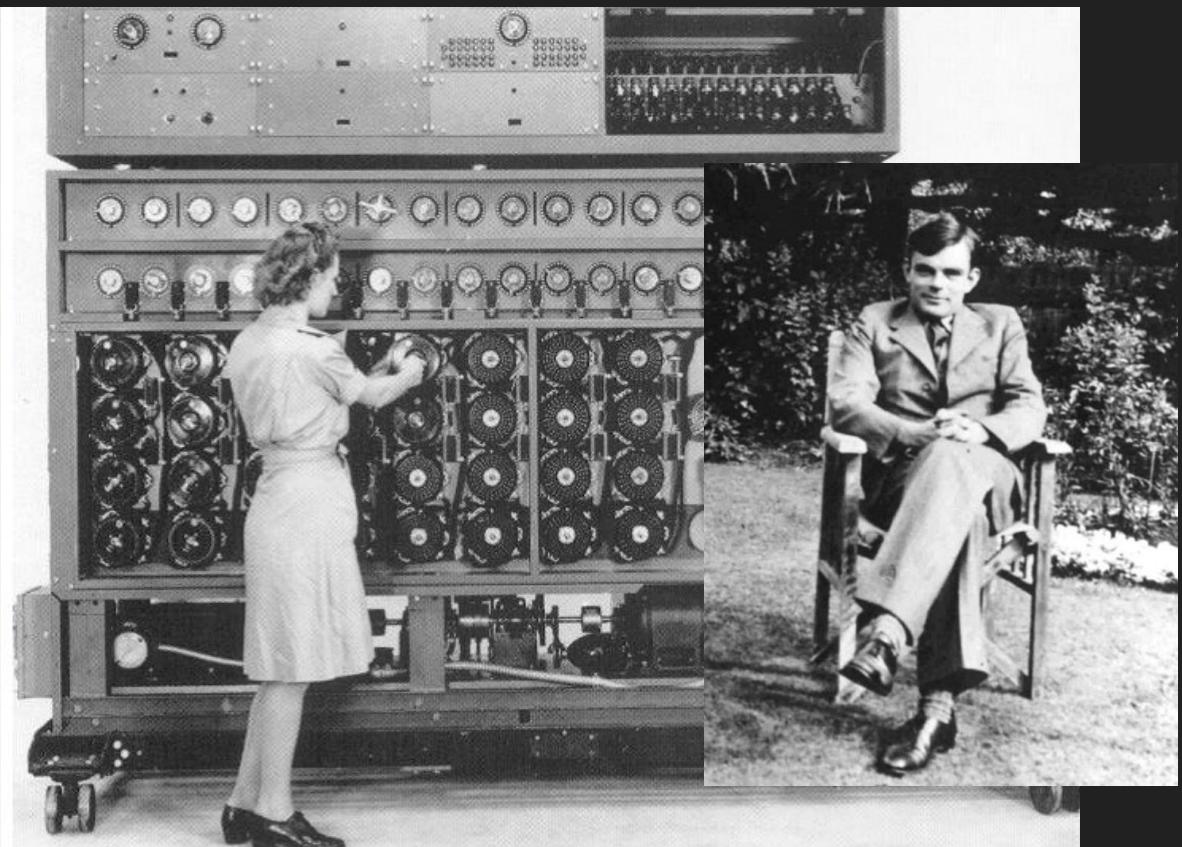
To play this game with the greatest chance of success, the experimenter cannot afford to exclude the possibility of any possible arrangement of soil fertilities, and his best strategy is to equalize the chance that any treatment shall fall on any plot by determining it by chance himself.

-R.A. Fisher



AJAX	K OF K	NITHSDALE	GREAT SCOTT	DUKE OF YORK
GREAT SCOTT	DUKE OF YORK	ARRAN COMRADE	IRON DUKE	EPICURE
IRON DUKE	EPICURE	AJAX	K OF K	NITHSDALE
K OF K	NITHSDALE	GREAT SCOTT	DUKE OF YORK	ARRAN COMRADE
UP TO DATE	KERR'S PINK	UP TO DATE	BRITISH QUEEN	
BRITISH QUEEN	TINWALD PERFECTION	EPICURE	KERR'S PINK	
KERR'S PINK	UP TO DATE	IRON DUKE	AJAX	
TINWALD PERFECTION	ARRAN COMRADE	BRITISH QUEEN	TINWALD PERFECTION	

World War 2: Turing and statistical cryptography



e.g., week 7 “women at the dawn...” (Abbate)

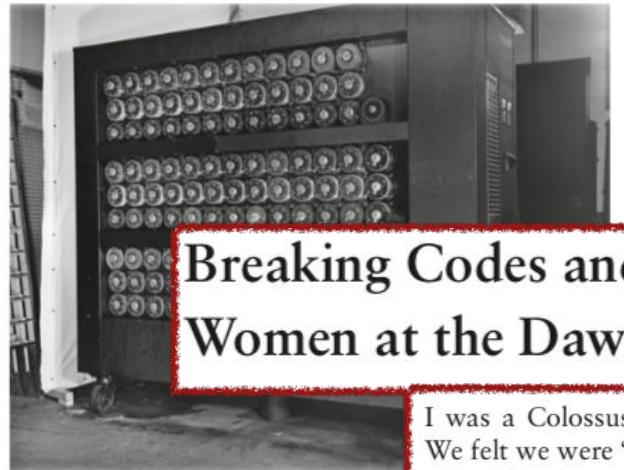
the theory



th

not d

sharon bertso



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

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

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 operator¹

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 programmer²

AI and its many winters

The Turing Test

1950: Alan Turing's "Computing Machinery and Intelligence"
(the "Turing Test")

A. M. Turing (1950) Computing Machinery and Intelligence. *Mind* 49: 433–460.

COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. The Imitation Game



Can machines think?

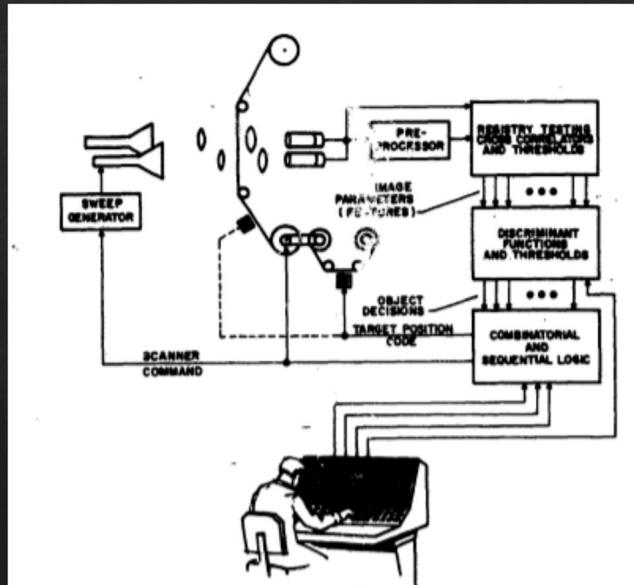
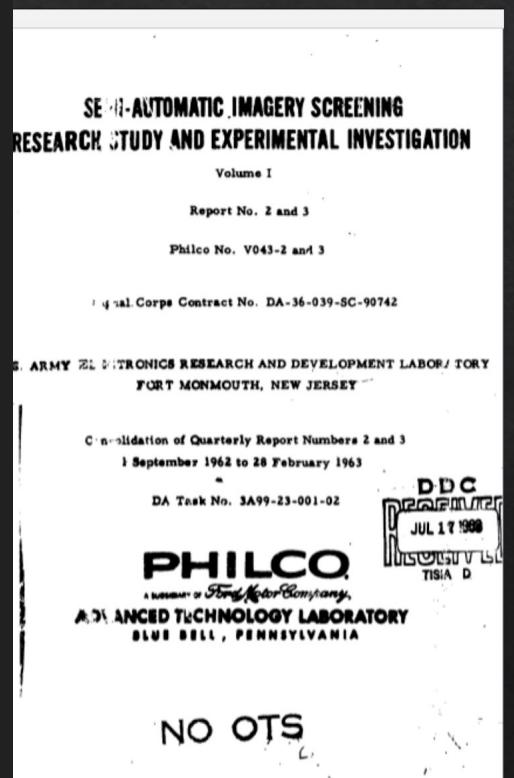
I propose to consider the question, "Can machines think?" This should begin with definitions of the meaning of the terms "machine" and "think." The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words "machine" and "think" are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, "Can machines think?" is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

BRACE YOURSELF!



AI WINTER IS COMING!!

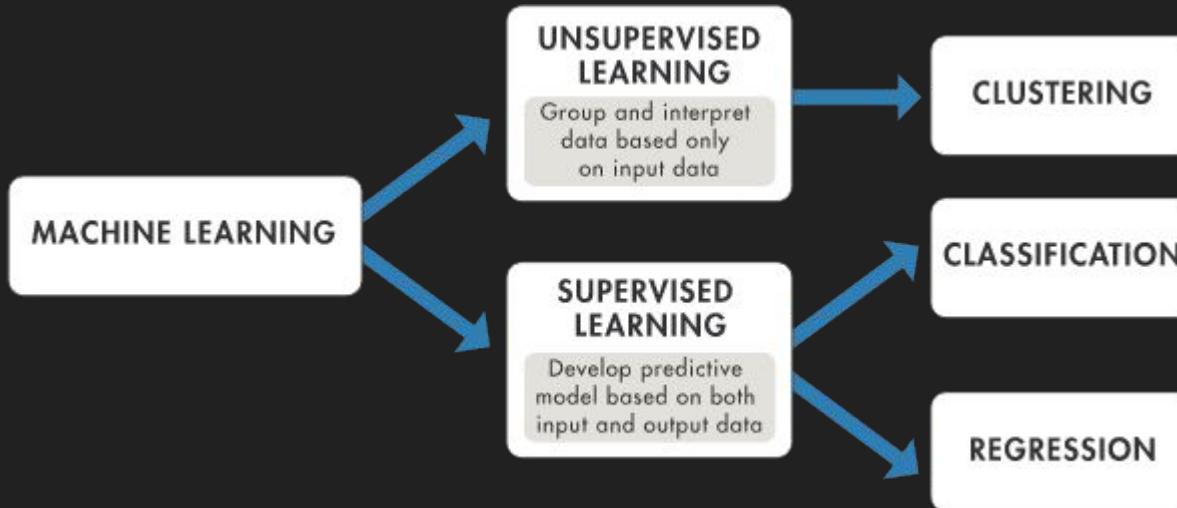
Pattern Recognition to . . .



Conceptual Block Diagram of Image Screening System



... Machine Learning



... and then to ethics & politics:

how did “ethics” meet “data”?

1. widely-condemned dumpster fires (1940s-1970s)
2. defining “practical” ethics (1970s)
3. enforcing / designing for ethics (1980s-present)

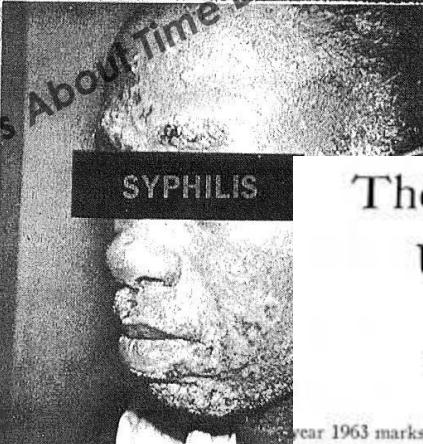
Research ethics (and the lack thereof)

**GERM WARFARE
DECLARED
AGAINST BLACKS!**

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

SEE ARTICLE INSIDE PAGE 2

Reprinted from the front page of the New York Times, July 26, 1963.



SYphilis

The Tuskegee Study of Untreated Syphilis

The 30th Year of Observation

DONALD H. ROCKWELL, MD; ANNE ROOF YOUNG, MD;
AND M. BRITTAINE MOORE, JR., MD, ATLANTA

year 1963 marks the 30th year of the evaluation of the effect of untreated syphilis in the male Negro conducted

tion such as this offered an opportunity to follow and study the disease over a long period of time. In 1932, a total

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 effective therapy was eventually developed.

The study was conducted to determine from autopsies what disease does to the human body.

Officials of the health service, who initiated the experiment, have long since retired. Recent 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 lack thereof)



"I was going to write an angry post about Facebook's emotional manipulation study, but then I got distracted by all the happy cat pictures they showed me."

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}

^aCore Data Science Team, Facebook, Inc., Menlo Park, CA 94025; and

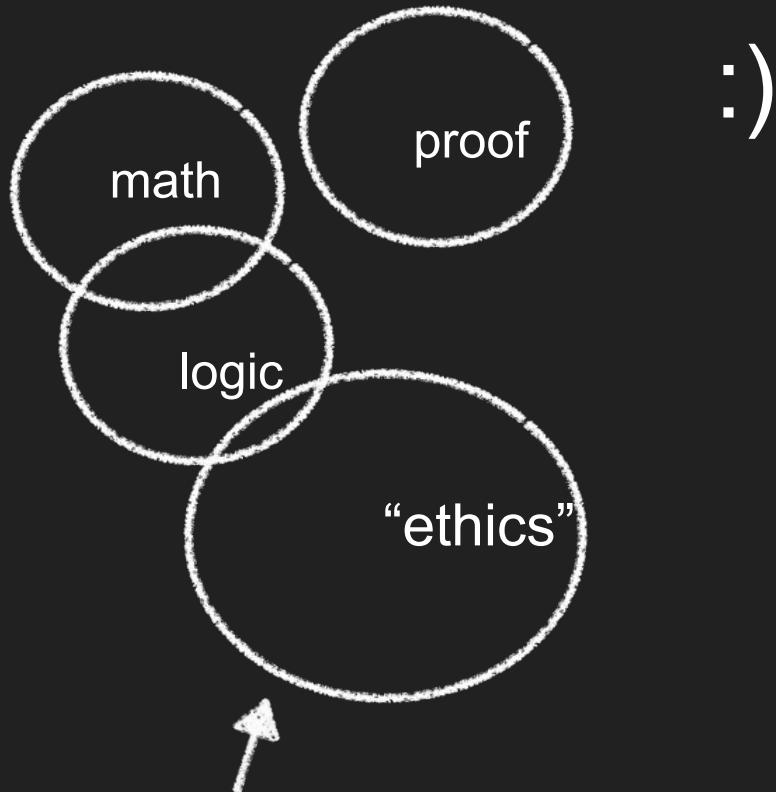
Departments of ^bCommunication and

^cInformation Science, Cornell University, Ithaca, NY 14853

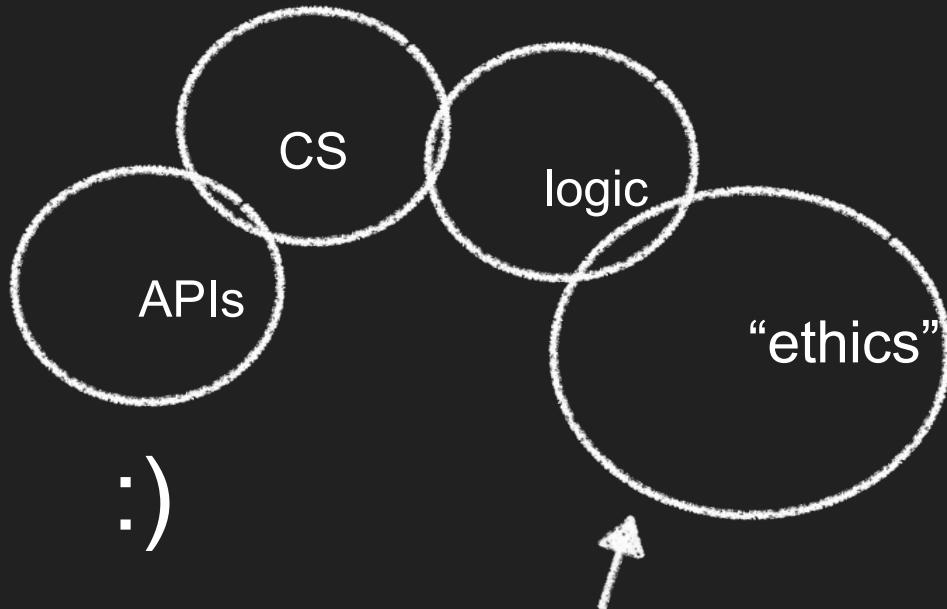
Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

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.

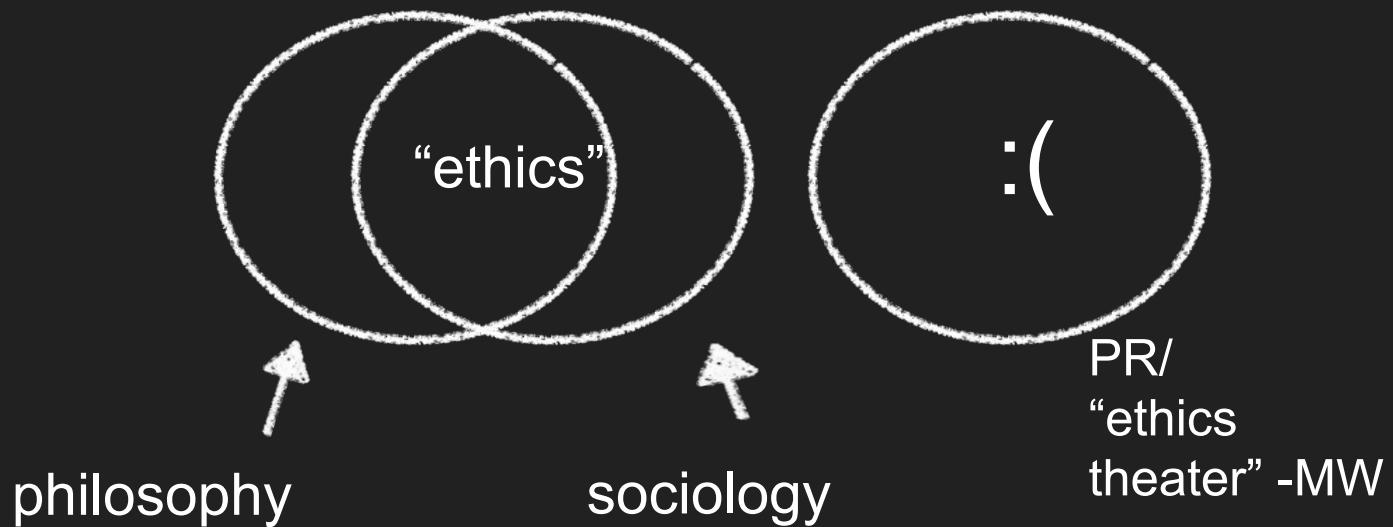


what we talk about when we talk about ethics

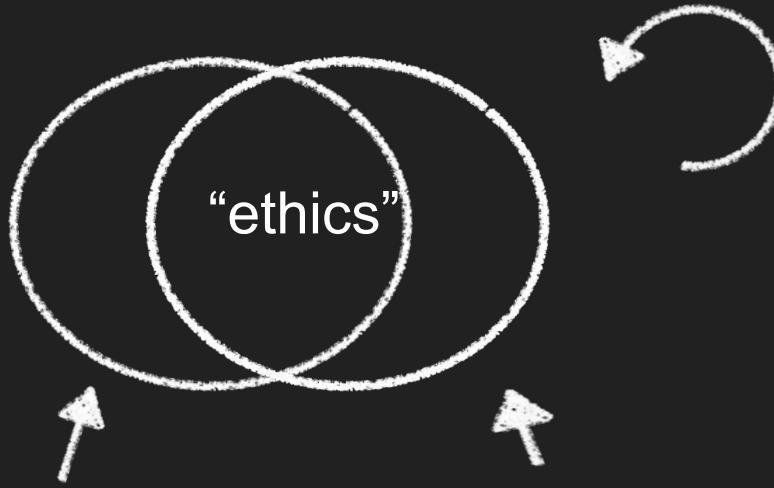


philosophy

what we talk about when we talk about ethics



what we talk about when we talk about ethics



philosophy

(define)

sociology

(design)

e.g., week 12 the ethics of data

Belmont principles (1978)

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

gives analytical, hierarchical, durable framework
for ethical audit of decisions,
from which rules, code, “design” should derive

Silicon Valley and the Attention Economy



Don't make the mistake of thinking you're Facebook's customer, you're not - you're the product,

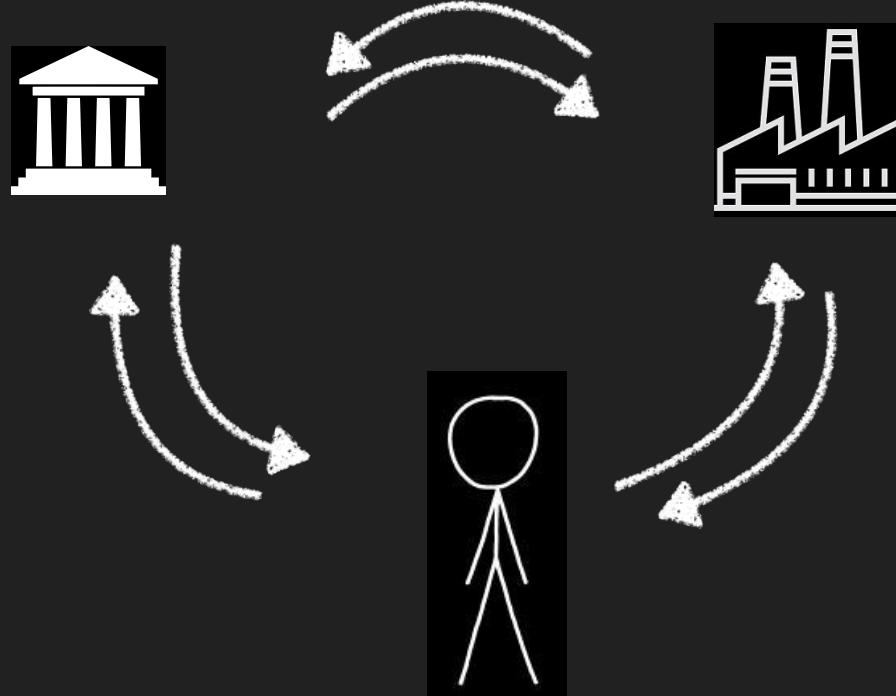
— Bruce Schneier —

AZ QUOTES

De-anonymization



e.g., week 14 (future) solutions



3-player unstable game (adapted from Janeway)

Weekly structure

Tuesday

Lecture and discussion

Expectation

arrive having done the week's readings

Thursday

Laboratory

Expectation

arrive with laptop ready to collaborate

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 readings
 - ...and presaging homeworks!
- Expectation
 - arrive with laptop ready to collaborate

“political” meaning impact on *human rights, justice, harms and benefits --- and on power*

No prerequisites

Instructors' permission
NOT
required

Required work

Postings on readings in Slack each week

Op-Ed (short but tight 750 words)

Participate in laboratory hours including posting to Slack as assigned during class

Problem sets (extensions of lab work)

Final paper

Two tracks

more technical background track (60%)

- pursue a semester long project culminating in a 15pp paper and any associated code
-
- complete 3 problem sets
-

more humanistic background track (60%)

- write a 10 pp paper on a topic of their choice
-
- complete 5 problem sets, these problem sets will involve both computational work 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.

Registration

History-APMA UN2901 section 001

Call Number 12059

- **The course satisfies one science requirement for CC and GS students and counts towards "nontech elective" credit for SEAS students.**
- **This course satisfies Barnard's T&D general education requirement.**
- **This course counts as a history course for history major.**
- **NO PREREQUISITES AND NO PROGRAMMING EXPERIENCE REQUIRED!**

Jupyter notebooks: code without any coding

Using

Codio platform

With everything you'll need installed

USING JUPYTER NOTEBOOKS

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 brains. You will undertake a series of statistical examinations AND you will also reflect critically on the nature of these data sets, and the lictiness of different computational operations on data sets.

```
In [ 1]: # usual preliminaries  
%matplotlib inline  
  
import pandas as pd  
import matplotlib.pyplot as plt  
import numpy as np  
plt.style.use('fivethirtyeight')  
plt.rcParams['figure.figsize'] = (15, 5)
```

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

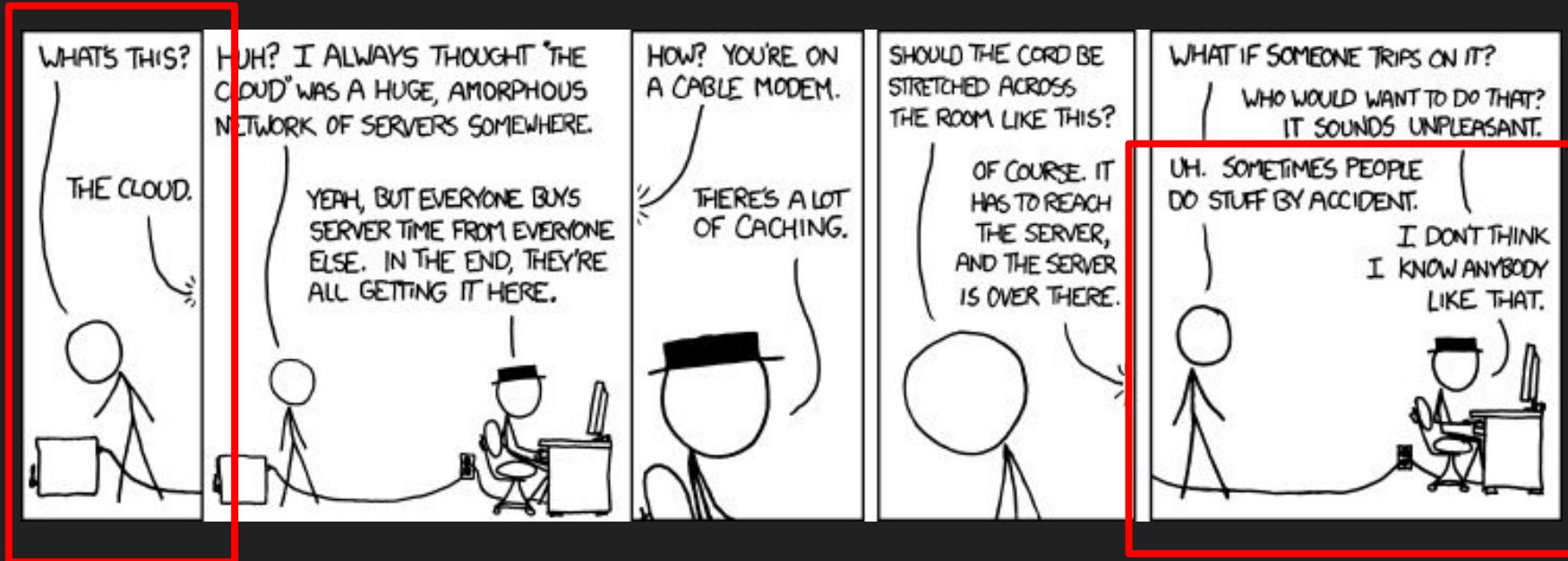
```
In [ 1]: data = pd.read_csv('http://www.scipy-lectures.org/_downloads/  
    ".*")
```

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

```
In [ 1]:
```

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

Codio: computing without computers



Codio: computing computer

The screenshot shows the Codio web interface. At the top, there's a navigation bar with links for Codio, Project, File, Edit, Find, View, Tools, Education, Help, Update..., Jupyter, Configuration, and a user profile for CWIGGINSS5. Below the navigation bar is a filetree sidebar titled "filetree" showing a directory structure under "WIGGINSS5". The main workspace contains a "README.md" file and a "Preview http..." tab. A Jupyter notebook titled "Lab01-20191024-UCI.ipynb" is open, showing its toolbar and content area.

Lab 1: Intro to Python & Data Provenance

Our lab today will have two parts: (1) a group activity about data sets and (2) an introduction to the use of the Jupyter notebook that we will use throughout the remainder of our course.

The Jupyter notebook is a helpful tool for programming in Python (a computer programming language). This lab will help you become more familiar with some of the basics.

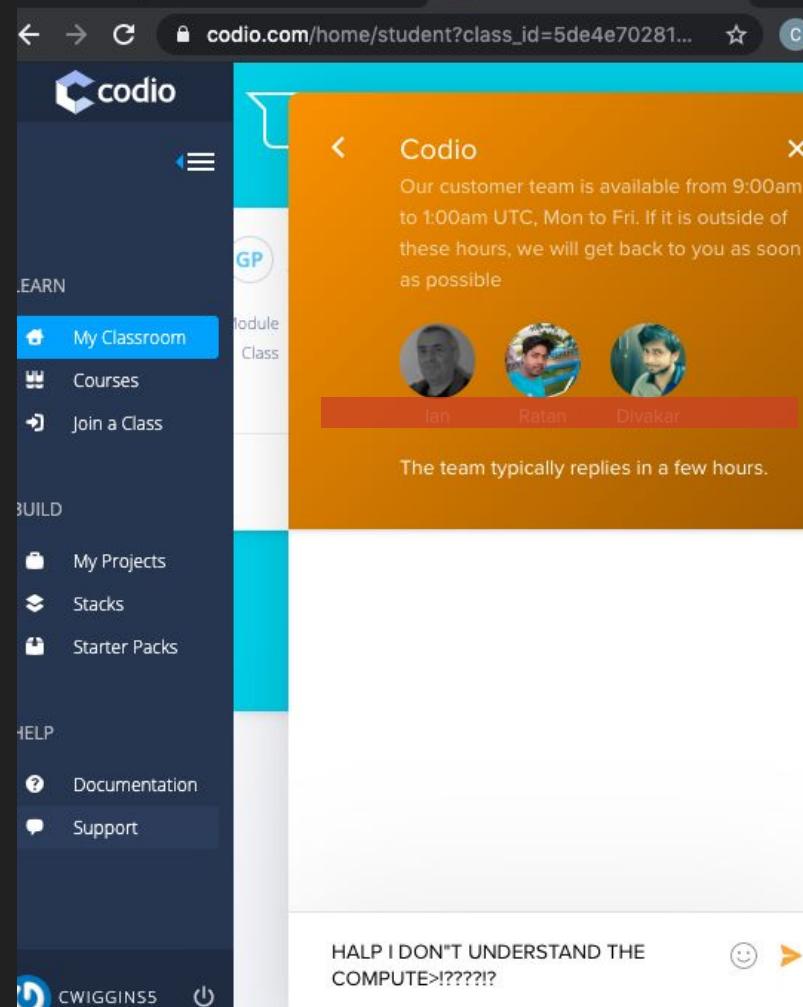
Very first steps with Jupyter

The following assumes you are using the Chrome web browser. (Using a different browser

To get to your notebook, we'll first need to run and connect to Jupyter. We'll be using Codio to do our work in Jupyter. To access Codio, go to the course website "powered by Canvas" via this [link](#), taking the login option to the left:

The screenshot shows the Courseworks@Columbia website at https://courseworks.columbia.edu/welcome/. The header includes the Columbia University logo and the text "COURSEWORKS@COLUMBIA POWERED BY CUIT". The footer provides links for Support for Students, Faculty, Administrators, Collaboration, and Copyright.

Codio: computing without computers



Slack-dom

#discussion

☆ | 57 | 0 | a chance for students to discuss by midnight on Saturday (preceding the M

Saturday, March 17th

researcher using data to predict future outcomes, 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.

✓ 1

0 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.

Search

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)

✓ 1 1 1

3 replies

chris wiggins 19 days ago
Yes I enjoyed that docu

This might be worth a sk in AI and creativity, this along musicians:
<https://muse.jhu.edu/art>

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

First readings for next Tuesday

danah boyd & Kate Crawford

CRITICAL QUESTIONS FOR BIG DATA

Provocations for a cultural,
technological, and scholarly
phenomenon



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

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