CSC2552 Topics in Computational Social Science: Al, Data, and Society

Lecture 1: Introduction to Computational Social Science Prof. Ashton Anderson, Fall 2023



	Week	Date	Торіс	Reviews Due	Textbook Readings
	1	9/7	Introduction to computational social science		Ch. 1
	2	9/14	Introduction to computational social science cont'd		Ch. 1
*	3	9/21	Observational studies 1	9/20 9:00pm	Ch. 2
\star	4	9/28	Observational studies 2	9/27 9:00pm	Ch. 2
\star	5	10/5	Experiments 1	10/4 9:00pm	Ch. 4
	6	10/12	Project proposals		
\star	7	10/19	Experiments 2	10/18 9:00pm	Ch. 4
\star	8	10/26	Asking questions	10/25 9:00pm	Ch. 3
\star	9	11/2	Deep learning	11/1 9:00pm	
\star	10	11/16	Ethics in computational social science	11/15 9:00pm	Ch. 6
	11	11/23	Project presentations (Part 1)		
	12	11/30	Project presentations (Part 2)		



Readymades

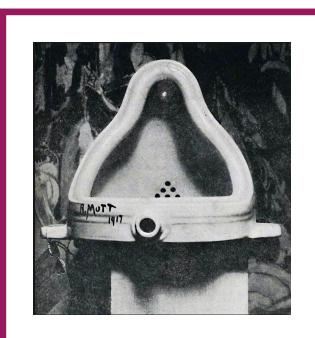
Computational social science in 7 easy pieces





Custommades

Ways of doing computational social science



Observational analyses

HumanNaturalFieldLabcomputationexperimentsSurveysexperimentsstudies

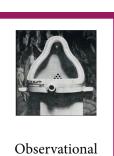


Observational analyses of existing data

- Massive datasets of all kinds of human behaviour are now available for study
 - Wikipedia, GPS traces, health databases, Facebook, Twitter, Reddit, reviews, purchases, dating, invitations, exercise apps, etc., etc...
- Key part of the "socioscope": huge traces of things that we couldn't see before
- Lack of detail/fidelity in individual records is hopefully made up for by large numbers of records (hope: small noisy errors cancel out, big patterns are signal)

"Big data" / "Found data"

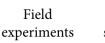




analyses

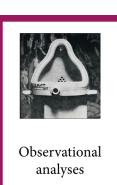
Human computation

Natural

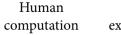


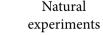
Ten common characteristics of big data

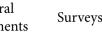
- **Big**: statistical power, rare events, fine resolution
- Always-on: unexpected events, real-time measurement
- Nonreactive: measurement probably won't change behaviour
- **Incomplete**: probably won't have the ideal information you want
- **Inaccessible**: difficult to access (gov't, companies)
- **Nonrepresentative**: bad out-of-sample generalization (good in-sample)
- **Drifting**: Population drift, usage drift, system drift
- Algorithmically confounded: want to study behaviour, not an algorithm
- **Dirty**: Junk, spam
- Sensitive: Private, hard to tell what's sensitive



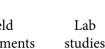






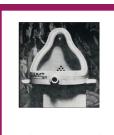






Observing Behaviour: Three research strategies

- 1. Counting things
- 2. Forecasting/nowcasting
- 3. Approximating experiments



Observational analyses

Human computation

Natural experiments

Surveys

Field experiments

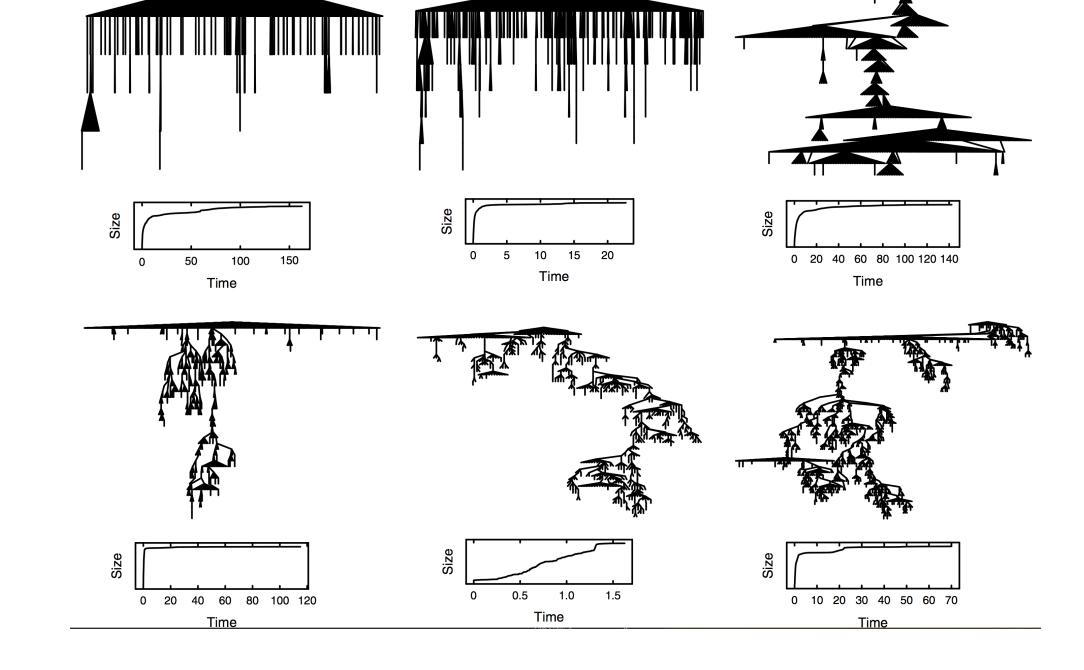


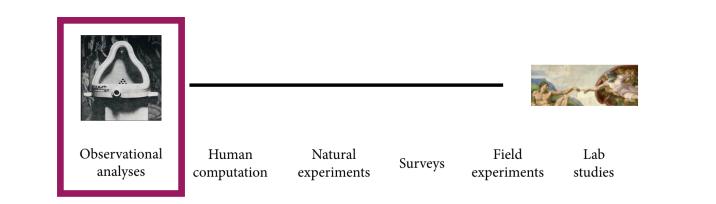
Lab studies

Observing Behaviour: 1. Counting Things

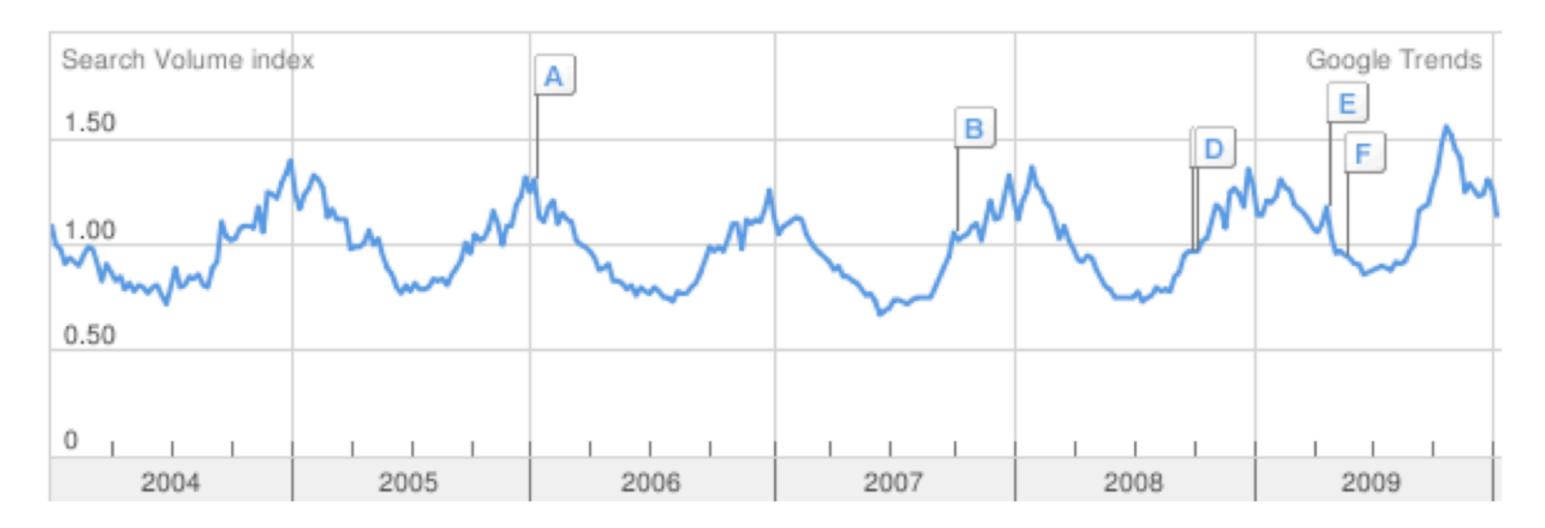
Example: Measuring viral vs. broadcast diffusion on Twitter

With newfound datasets and computational resources, many valuable initial contributions are measurements of quantities we couldn't measure before \rightarrow counting at scale



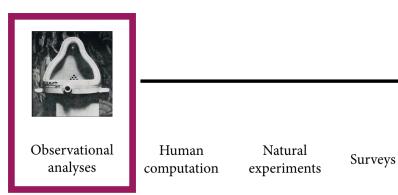


Google Flu Trends



Search volume for the term "cough"

Idea: find 50 most correlated search query volume trends with flu data



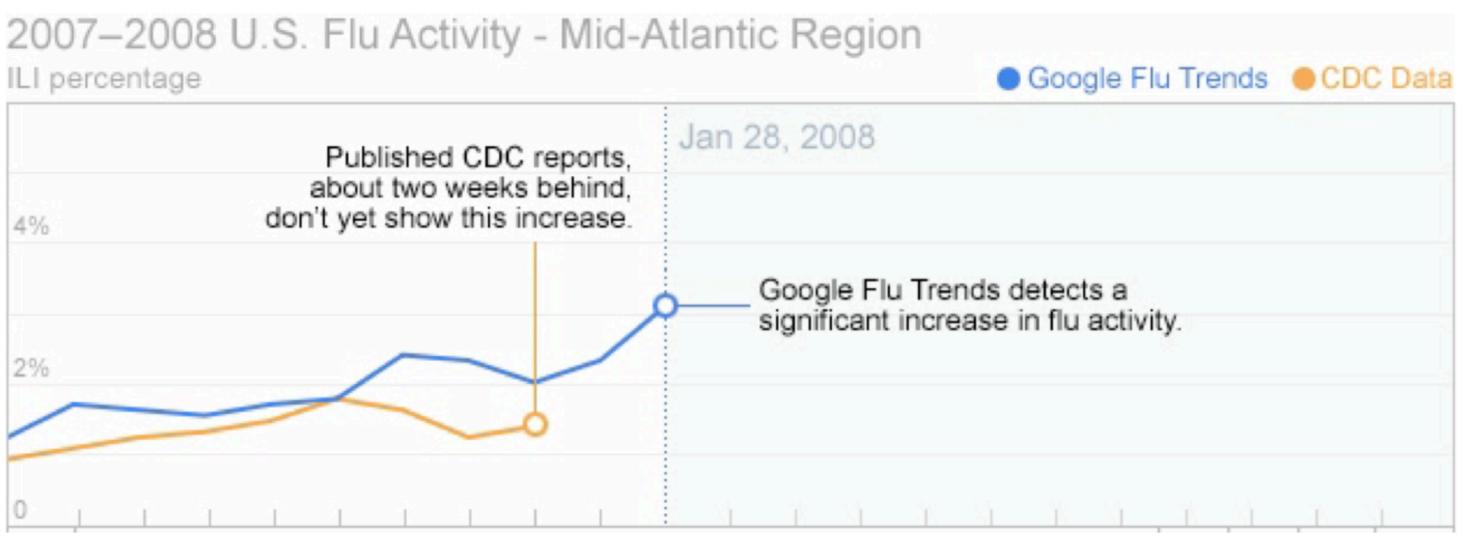
Lab

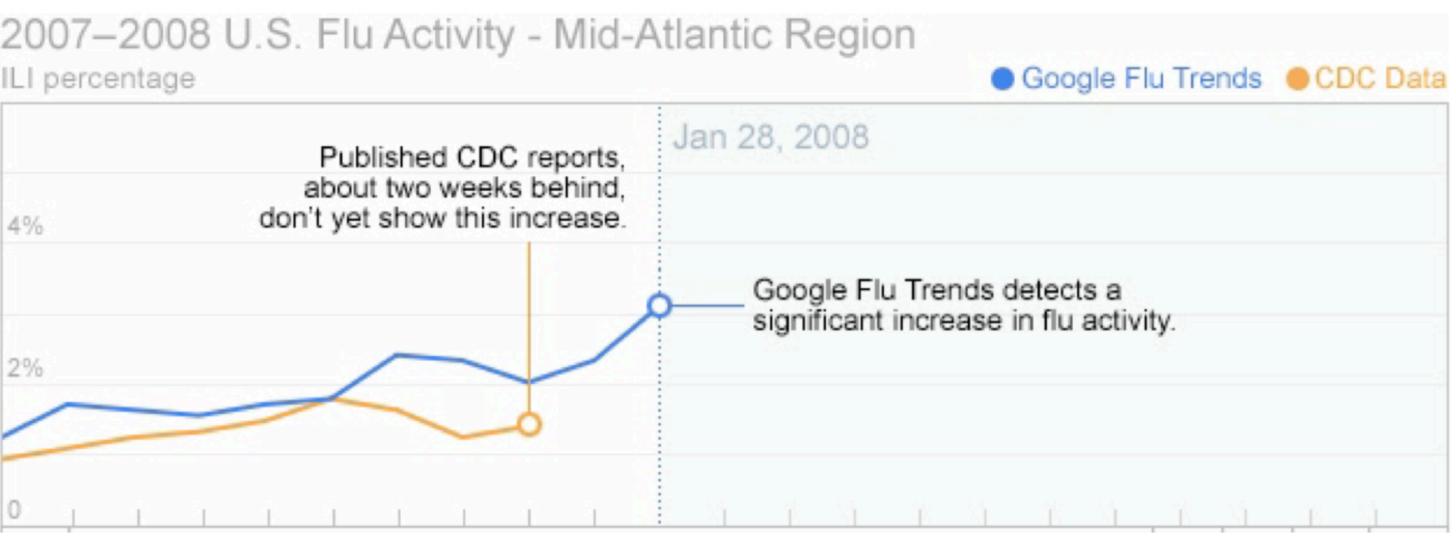
studies

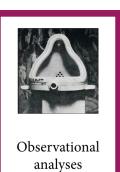
Field

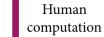
experiments

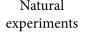
The flu has a 1-2 week lag from when cases are reported to when the CDC releases official stats



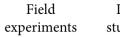








Surveys



Lab studies

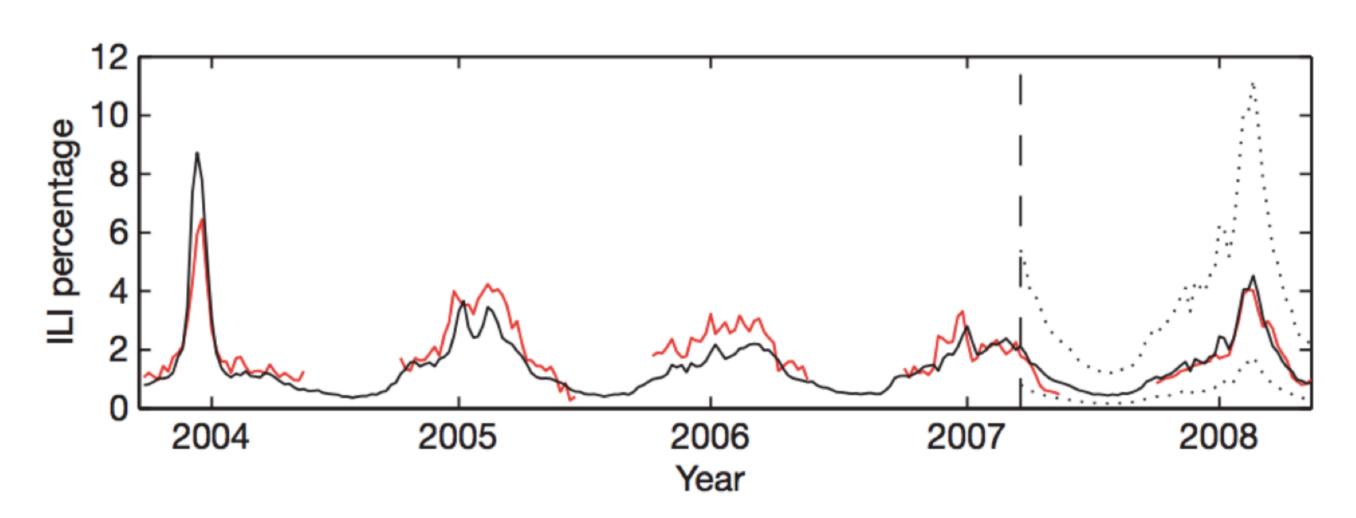
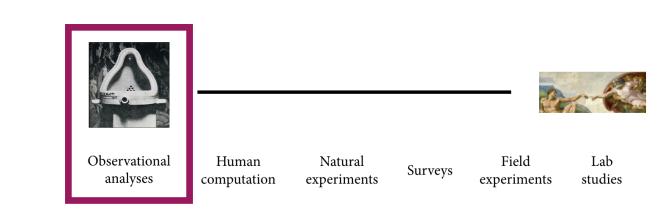
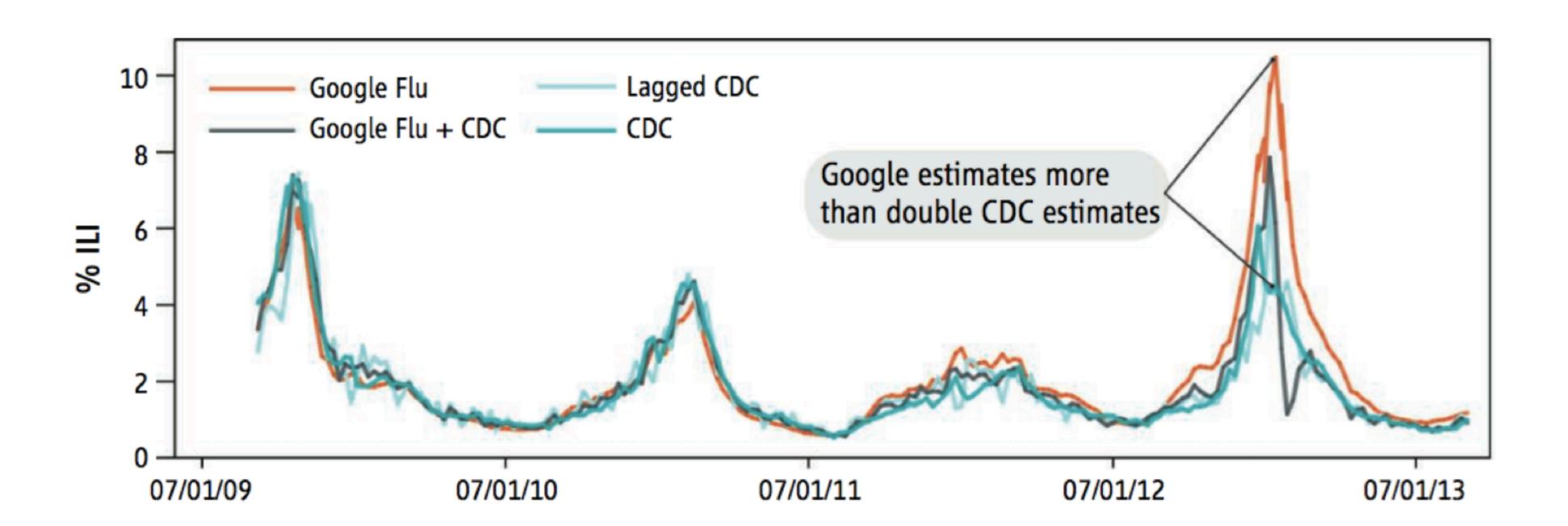
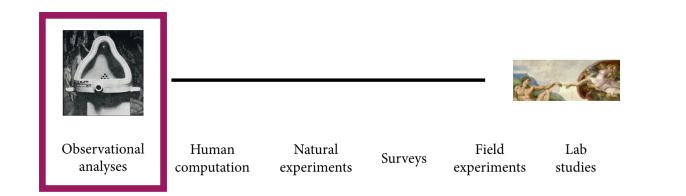


Figure 2 | A comparison of model estimates for the mid-Atlantic region (black) against CDC-reported ILI percentages (red), including points over which the model was fit and validated. A correlation of 0.85 was obtained over 128 points from this region to which the model was fit, whereas a correlation of 0.96 was obtained over 42 validation points. Dotted lines indicate 95% prediction intervals. The region comprises New York, New Jersey and Pennsylvania.

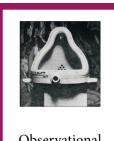




Soon after Google Flu Trends launched, it was drastically off



Media attention "Bird flu", "swine flu" Algorithm changes Starting suggesting search terms "Social hacking" (Trolling) Hey look we can screw up Google's flu predictions



Observational analyses

Human computation

Natural experiments

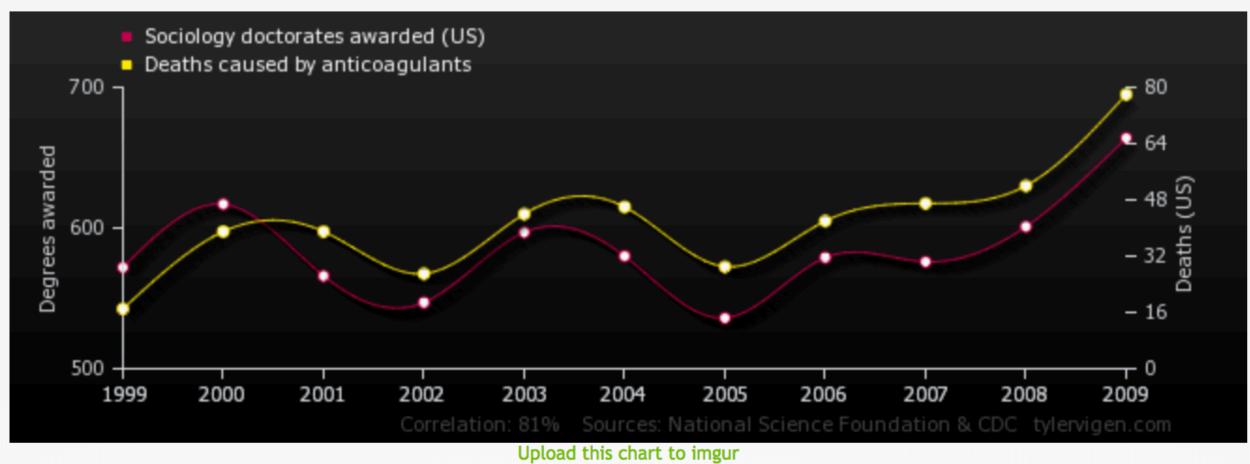
Surveys

Field experiments

Lab studies



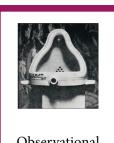
Sociology doctorates awarded (US) correlates with **Deaths caused by anticoagulants**



	<u>1999</u>	<u>2000</u>	<u>2001</u>	<u>2002</u>	<u>2003</u>	<u>2004</u>	<u>2005</u>	<u>2006</u>	<u>2007</u>	<u>2008</u>	<u>2009</u>
Sociology doctorates awarded (US) Degrees awarded (National Science Foundation)											664
Deaths caused by anticoagulants Deaths (US) (CDC)	17	39	39	27	44	46	29	42	47	52	78

Correlation: 0.811086

Correlation and causation



Observational analyses

Human computation

Natural experiments

Surveys

Field Lab studies



experiments

People who died by falling out of their bed correlates with Lawyers in Puerto Rico



Correlation and causation

<u>:001</u>	<u>2002</u>	<u>2003</u>	<u>2004</u>	<u>2005</u>	<u>2006</u>	<u>2007</u>	<u>2008</u>	<u>2009</u>	<u>2010</u>
516	551	59 4	503	621	626	690	737	780	718
1,071	10,947	11,209	11,191	11,805	11,767	12,142	12,454	13,071	13,282



Observational analyses

Human computation

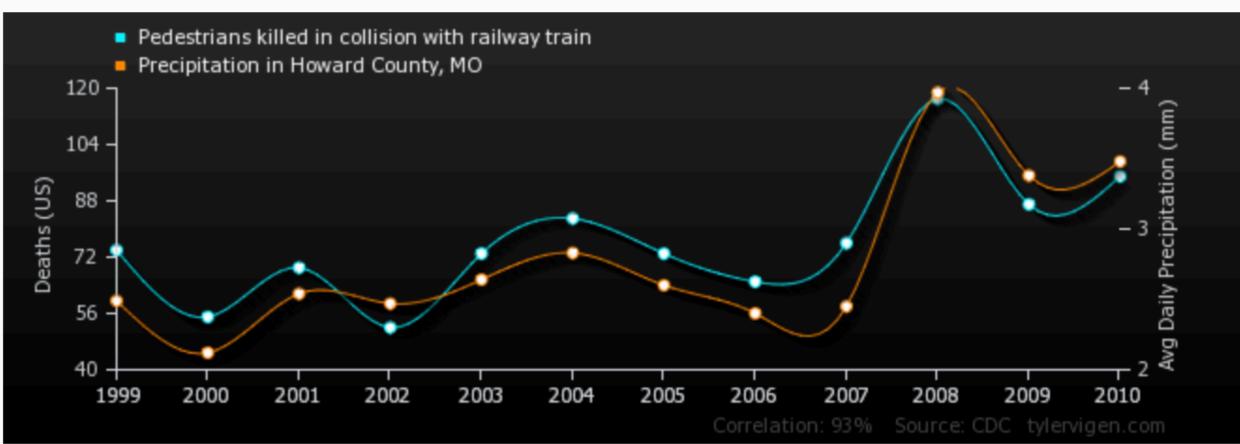
Natural experiments

Surveys

Field Lab studies experiments



Pedestrians killed in collision with railway train correlates with **Precipitation in Howard County, MO**

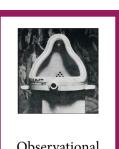


	<u>1999</u>	<u>2000</u>	<u>2001</u>	<u>2002</u>	<u>2003</u>	<u>2004</u>	<u>2005</u>	<u>2006</u>	<u>2007</u>	<u>2008</u>	<u>2009</u>	<u>2010</u>
Pedestrians killed in collision with railway train Deaths (US) (CDC)	74	55	69	52	73	83	73	65	76	117	87	95
Precipitation in Howard County, MO Avg Daily Precipitation (mm) (CDC)	2.49	2.12	2.54	2.47	2.64	2.83	2.6	2.4	2.45	3.97	3.38	3.48

Correlation: 0.92783

Correlation and causation

Upload this chart to imgur



Observational analyses

Human computation

Natural experiments

Surveys

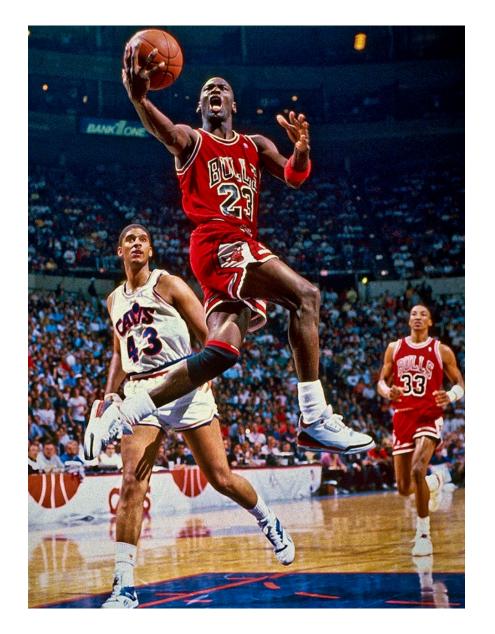


Field experiments

Lab studies

Perils of big data

"When you have large amounts of data, your appetite for hypotheses tends to get even larger. And if it's growing faster than the statistical strength of the data, then many of your inferences are likely to be false. They are likely to be white noise." — Michael Jordan





Observational analyses

Human computation

Natura experiments

Surveys

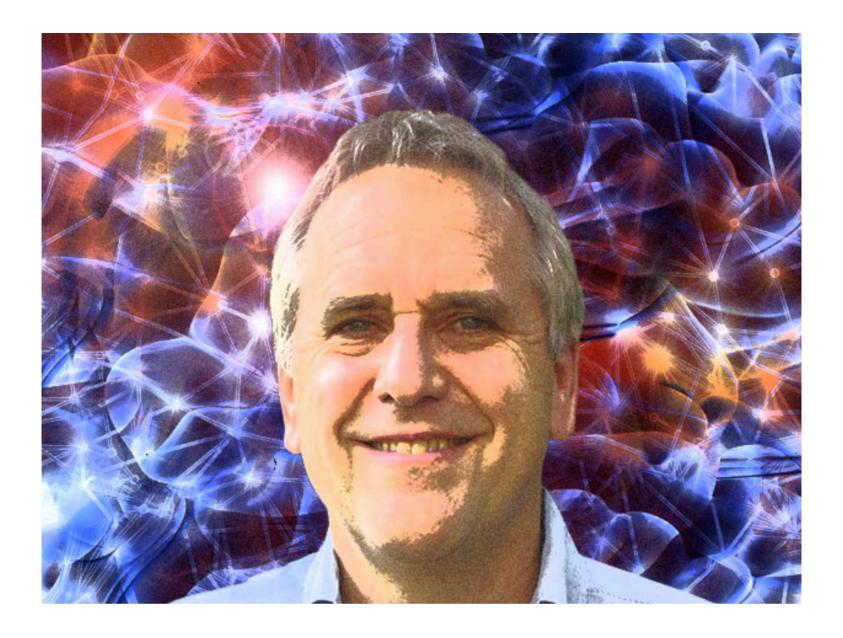
Field experiments

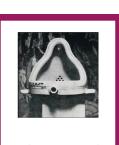




Perils of big data

"When you have large amounts of data, your appetite for hypotheses tends to get even larger. And if it's growing faster than the statistical strength of the data, then many of your inferences are likely to be false. They are likely to be white noise." — Michael Jordan





Observational analyses

Human computation

Natural experiments

experiments





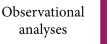
Observing Behaviour: 3. Approximating Experiments

Some clever strategies allow us to do "causal inference": make causal claims from observational data (i.e. arrive at experiment-like conclusions without actually running an experiment)

One well-known technique is instrumental variables: exploit natural variation in something to make a causal claim

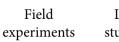
Rain \rightarrow Exercise Friends exercising → You exercise?





Human computation

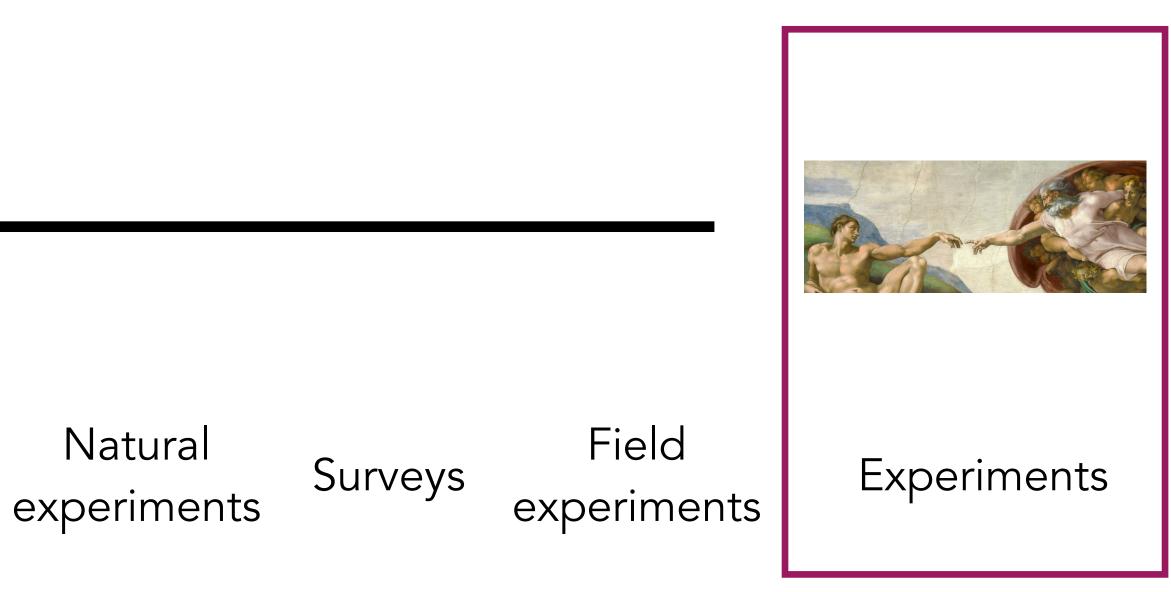




Ways of doing computational social science



Observational Human Nation analyses computation experience



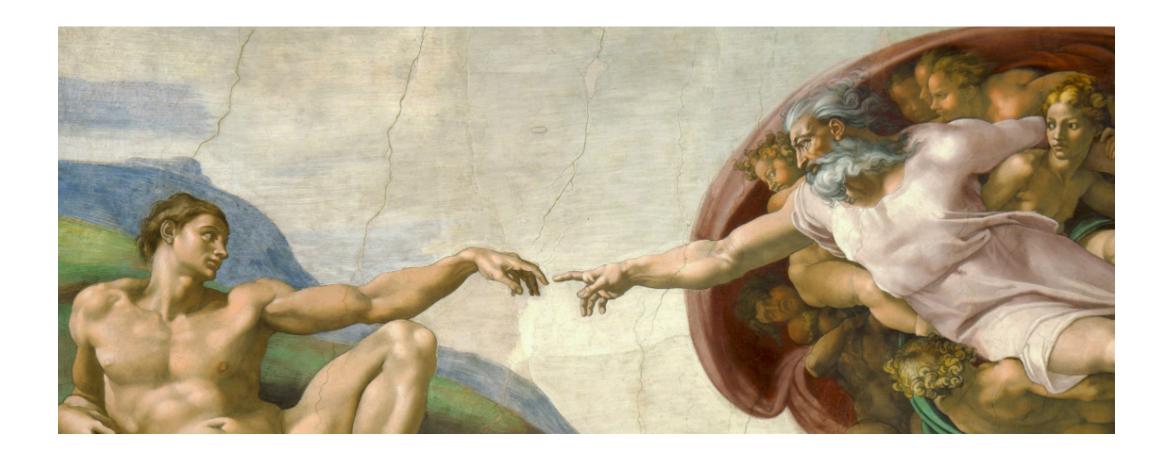
Experiments

On the other end of the spectrum is experimentation

The goal is to learn about causal relationships (cause-and-effect questions)

The strategy is to directly manipulate the environment and observe the consequences

Design the ideal scenario that will create just the data you need to answer your question



Experiments

Here, researchers intervene in the world to isolate and study a specific question

Nomenclature:

"Experiment": perturb and observe "Randomized controlled experiment": Intervene for one group, don't for another (randomly)

Correlation is not causation

Observational data often riddled by unknown or hard-to-control confounding variables

E.g. Do students learn more in schools that offer high teacher salaries? What's an observational way to study this question? What's wrong with it? What's an experimental way to study this question?

What's wrong with it?





Observational analyses

Human computation

Natural experiments

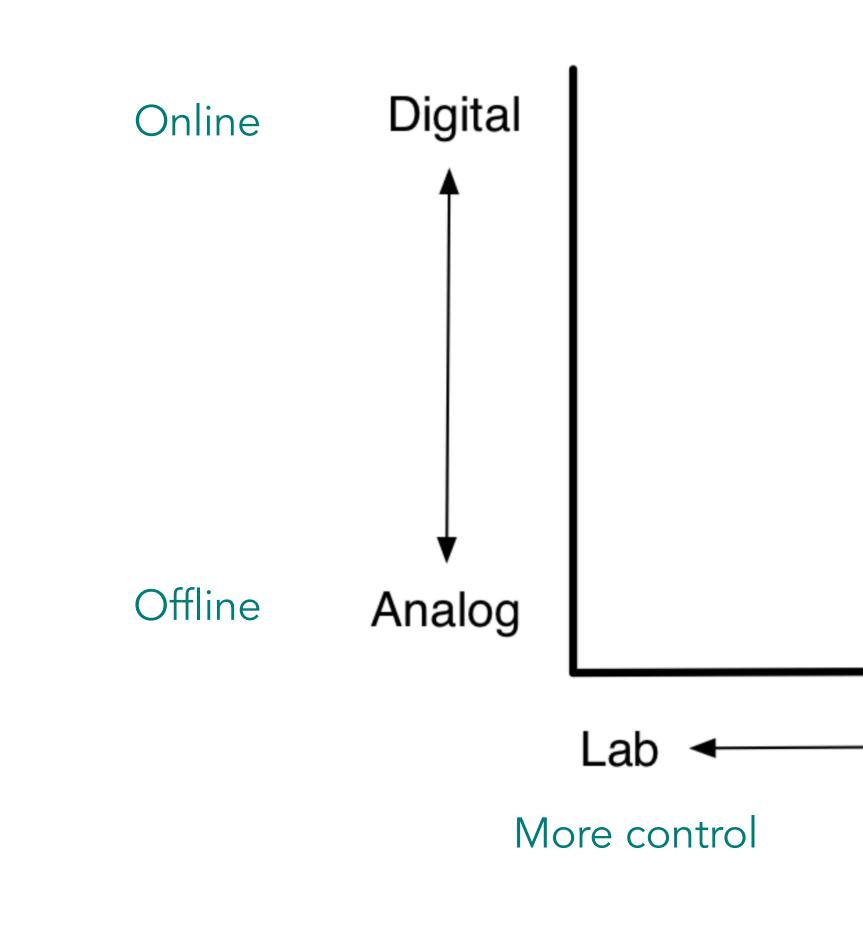
Surveys

experin ents

Lab studies



Experiments



► Field

More real





Observational analyses

Human

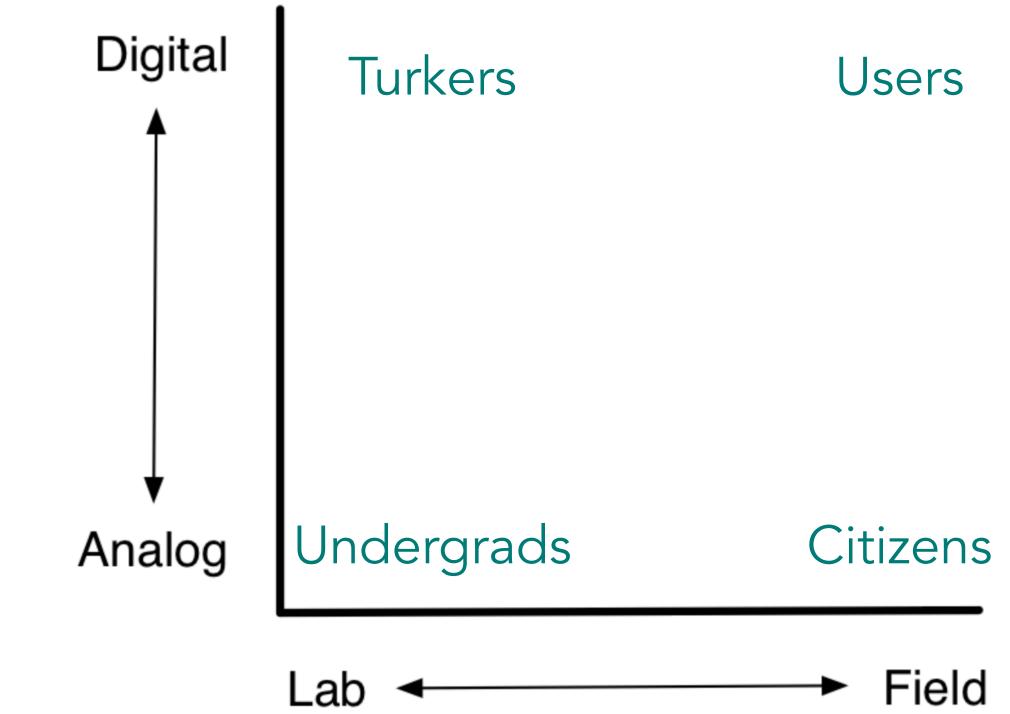
Natural computation experiments

Surveys

experin

Lab studies nts

Con Con









Observational analyses

Human computation

Natural experiments

Surveys

experin

Lab studies nts

Three major components of rich experiments

- 1. Validity
- 2. Heterogeneity
- 3. Mechanisms





Observational analyses

Human computation

Natural experiments

Lab studies

Survey

experi

Three major components of rich experiments: 1. Validity

Validity: How general are the results?

Types of validity:

- 1. Statistical conclusion validity: were the stats done right?
- 2. Internal validity: was the experiment done right?
- 3. Construct validity: are we measuring the right thing?
- 4. External validity: is this applicable in other settings?





Observational analyses

Human computation

Natural experiments





Three major components of rich experiments: 2. Heterogeneity

Barebones experiment: measure the average treatment effect (ATE)

But in social research, people almost always vary.

Digital research presents many more opportunities to measure how causes affect people differently





Observational analyses

Human computation

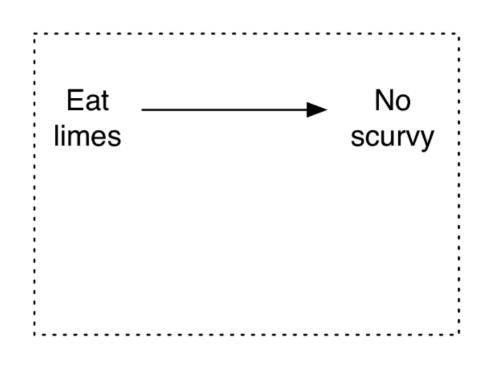
Natural experiments



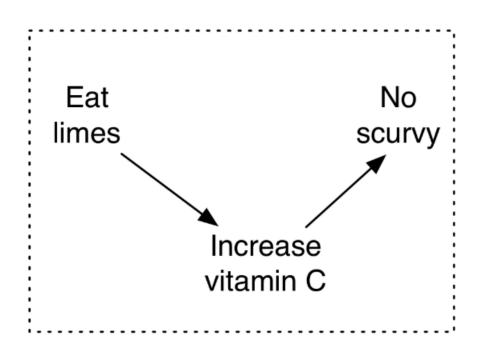
Three major components of rich experiments: 3. Mechanisms

Barebones experiment: measure what happened.

Mechanisms: why and how did it happen?



Causal effect without mechanism



Causal effect with mechanism







Observational analyses

Human computation

Natural Su experiments

Surveys exper

nts studies

Ways of doing computational social science



Observational analyses

Human computation

Natural experiments



ural Surveys nents ex

Field experiments

Experiments

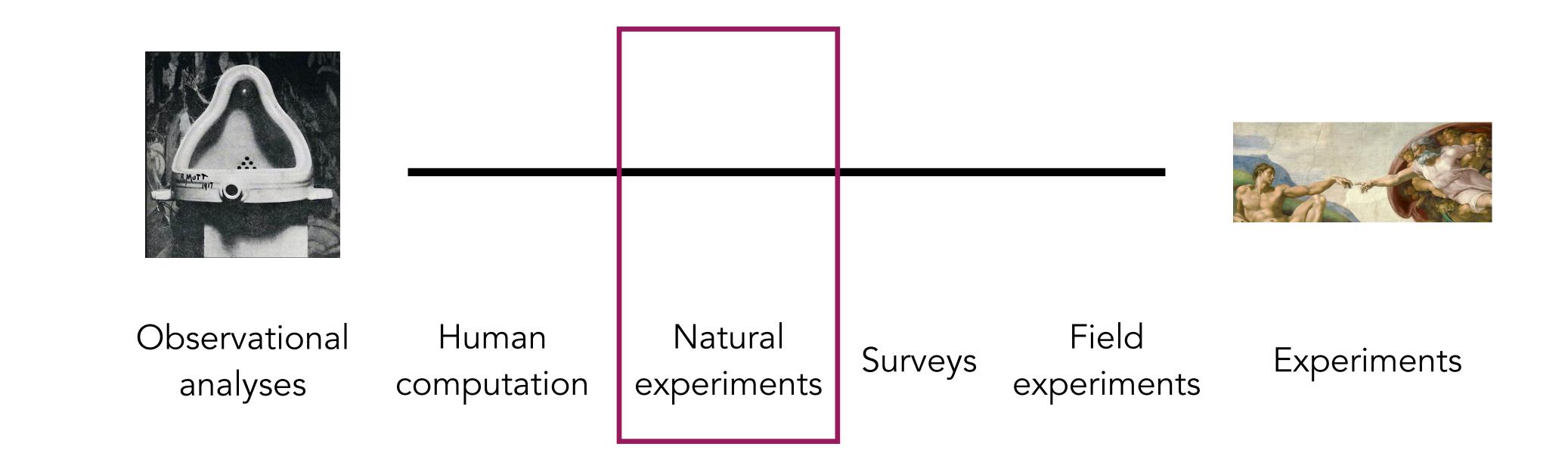
Human computation

- Online crowdsourcing platforms allow dividing work into microtasks
- resources (Wikipedia etc.)

Human-in-the-loop computing, modern-day lab studies, mass collaboration to build big

Artificial Artificial Intelligence	Your Account		Diet ,700 HITs ilable now	mar Hafner Account	Settings Sign Out Hel _l
Find HITs v containing	All HITS HITS A V	railable To You HITs Assigned To You that pay at least \$		you are qualified aster Qualification 60)
All HITS 1-10 of 2317 Results Sort by: HIT Creation Date (newest first) T	Show all details	Hide all details			1 <u>2 3 4 5</u> → <u>Next</u> ≫ <u>Last</u>
CTRP: Type name, date and total of a receipt			Requ	est Qualification (Why?)	View a HIT in this group
Requester: CopyText Inc.	HIT Expiration Date:	Jul 10, 2015 (9 minutes 52 seconds)	Reward:	\$0.01	
	Time Allotted:	4 minutes	HITs Available:	35	ļ
Where are you? (2 second HIT) USA			Not Qualified to v	work on this HIT (Why?)	View a HIT in this group
Requester: techlist	HIT Expiration Date:	Jul 10, 2015 (9 minutes 52 seconds)	Reward:	\$0.02	
	Time Allotted:	1 minute 30 seconds	HITs Available:	1067	ļ
Where are you? (2 second HIT) Not USA or India					View a HIT in this group
Requester: <u>techlist</u>	HIT Expiration Date:	Jul 10, 2015 (9 minutes 52 seconds)	Reward:	\$0.02	
	Time Allotted:	1 minute 30 seconds	HITs Available:	1073	
Where are you? (2 second HIT) India			Not Qualified to v	work on this HIT (Why?)	<u>View a HIT in this group</u>
Requester: <u>techlist</u>	HIT Expiration Date:	Jul 10, 2015 (9 minutes 51 seconds)	Reward:	\$0.02	
	Time Allotted:	1 minute 30 seconds	HITs Available:	1071	
QC Reject - \$0.20 per media minute			Requ	est Qualification (Why?)	View a HIT in this group
Requester: Crowdsurf Support	HIT Expiration Date:	Jul 8, 2016 (51 weeks 6 days)	Reward:	\$0.20	
	Time Allotted:	6 hours	HITs Available:	7	
Find the count of comments on a website					View a HIT in this group
Requester: SDG Production	HIT Expiration Date:	Jul 13, 2015 (2 days 23 hours)	Reward:	\$0.02	
	Time Allotted:	10 minutes	HITs Available:	1	ļ
Classify Receipt			Not Qualified to v	work on this HIT (Why?)	<u>View a HIT in this group</u>
Requester: Jon Brelig	HIT Expiration Date:	Jul 17, 2015 (6 days 23 hours)	Reward:	\$0.02	
	Time Allotted:	20 minutes	HITs Available:	7948	

Ways of doing computational social science



Natural experiments

Sometimes observational data has some random component you can exploit, and analyze as a "natural" experiment

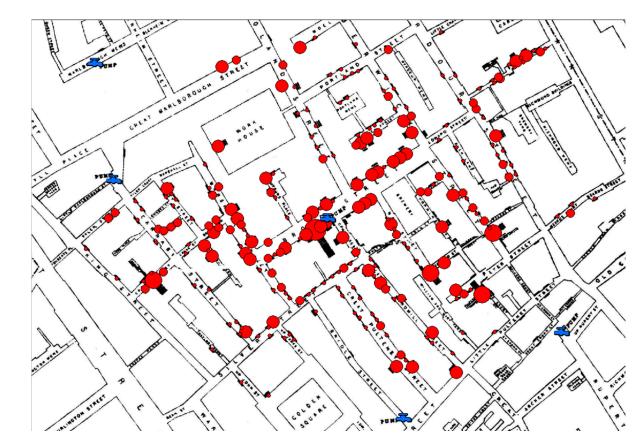


Cholera outbreak in London in 1850s

Natural experiments

- Physician John Snow produced a map suggesting particular water was the culprit
- Two main water suppliers: one from downstream Thames where raw sewage was dumped in the water (high attack rates), and one from upstream (low attack rates)
- Which supplier you had was arbitrary (varied even within same house, same neighbourhood, etc.)
- Exposure to polluted water was as-if random

Now: in large datasets, more opportunities to identify and argue for as-if random assignment



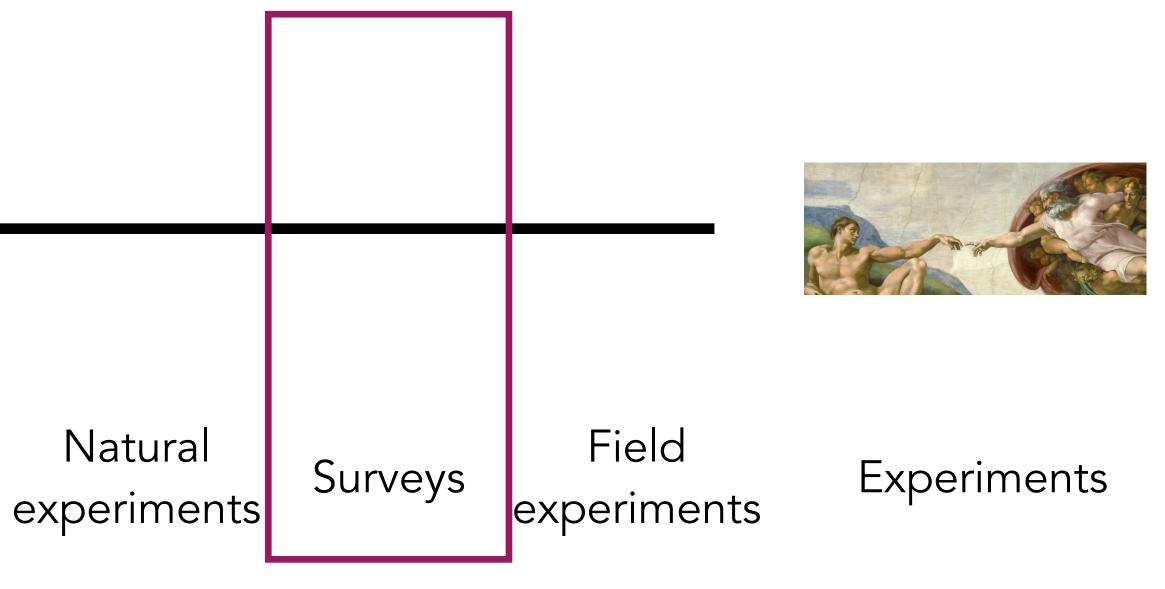
Cholera outbreak in London in 1850s



Ways of doing computational social science



Observational Human Nation analyses computation experience

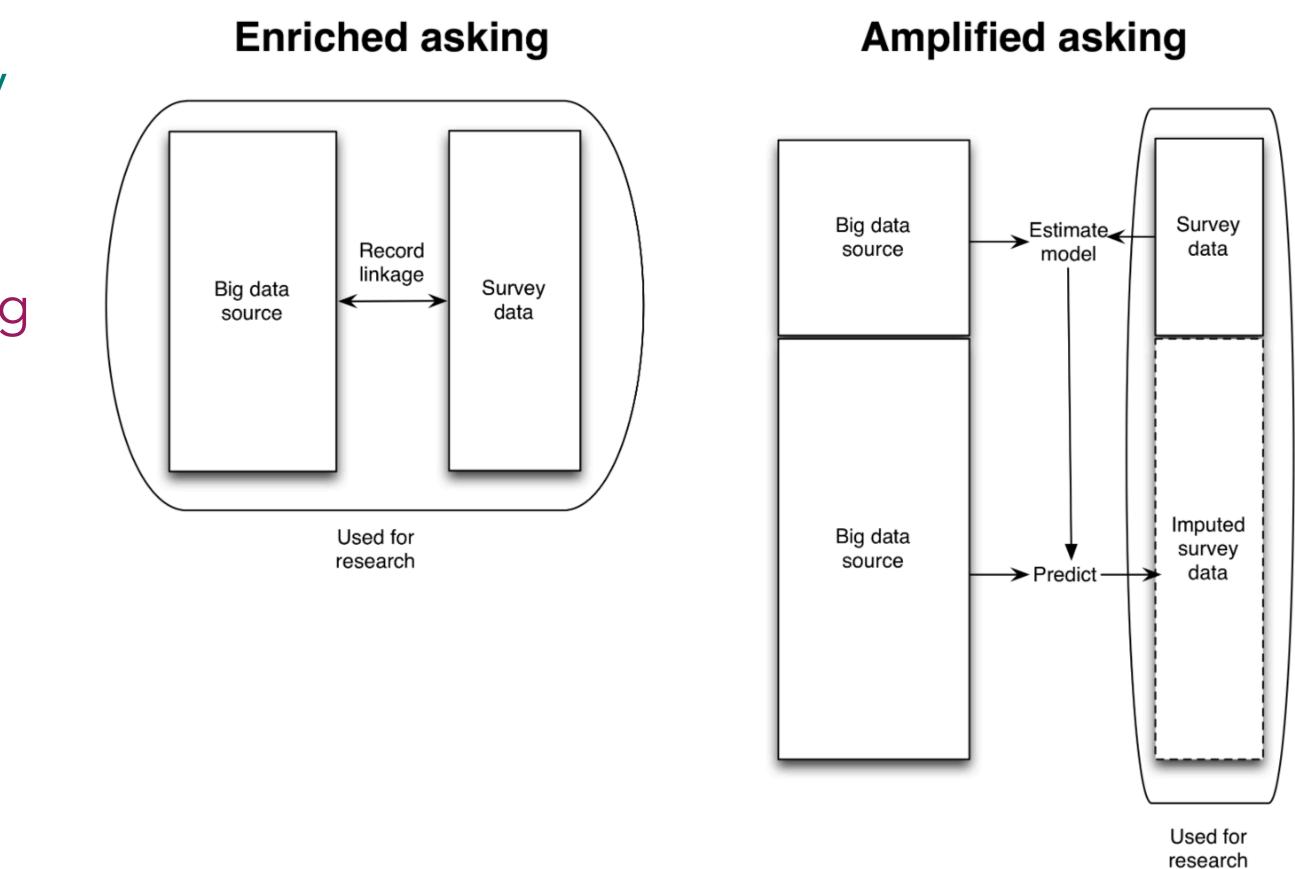


Surveys: asking questions

Social research has a unique advantage: we can ask our subjects what they're thinking!

Still the best way to learn the answer to many questions

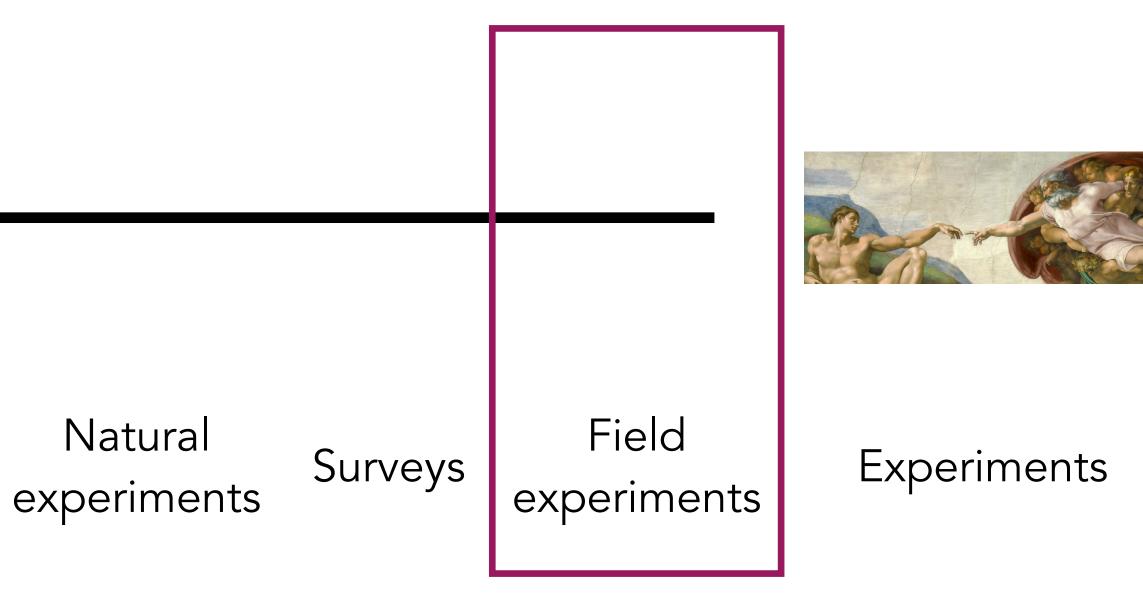
In the digital era, there are new ways of asking questions



Ways of doing computational social science



Observational Human Nation analyses computation experience

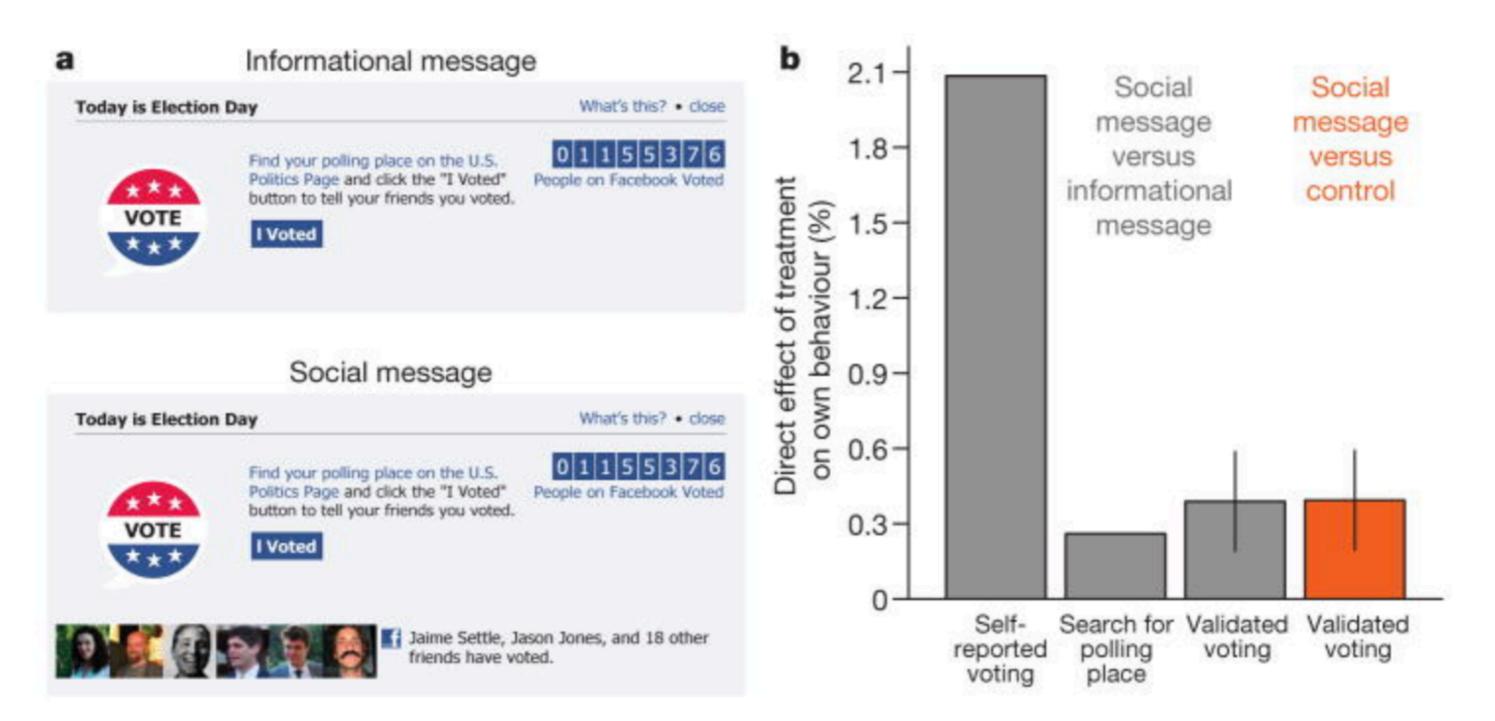


Field experiments

- Introducing a treatment into a real system
- Much more possible now with algorithmic systems

Voting experiment on Facebook

Figure 1



The experiment and direct effects

a, **b**, Examples of the informational message and social message Facebook treatments (**a**) and their direct effect on voting behaviour (**b**). Vertical lines indicate s.e.m. (they are too small to be seen for the first two bars).

~300,000 more validated votes

Al & Society: Algorithmic decision-making

St. George's Hospital in the UK developed an algorithm to sort medical school applicants. Algorithm trained to mimic past admissions decisions made by humans.

But past decisions were biased against women and minorities. It codified discrimination.

Ads by Google

We Found:Kristen Haring

 Contact Kristen Haring - Free Info! 2) Current Phone, Address & More.

www.peoplesmart.com/Kristen

Search by Phone Search by Email Background Checks Search by Address Public Records Criminal Records

Kristen Haring

Public Records Found For: Kristen Haring. Search Now. www.publicrecords.com/

Web search ads for "Kristen Haring"

Web search ads for "Latanya Farrell"

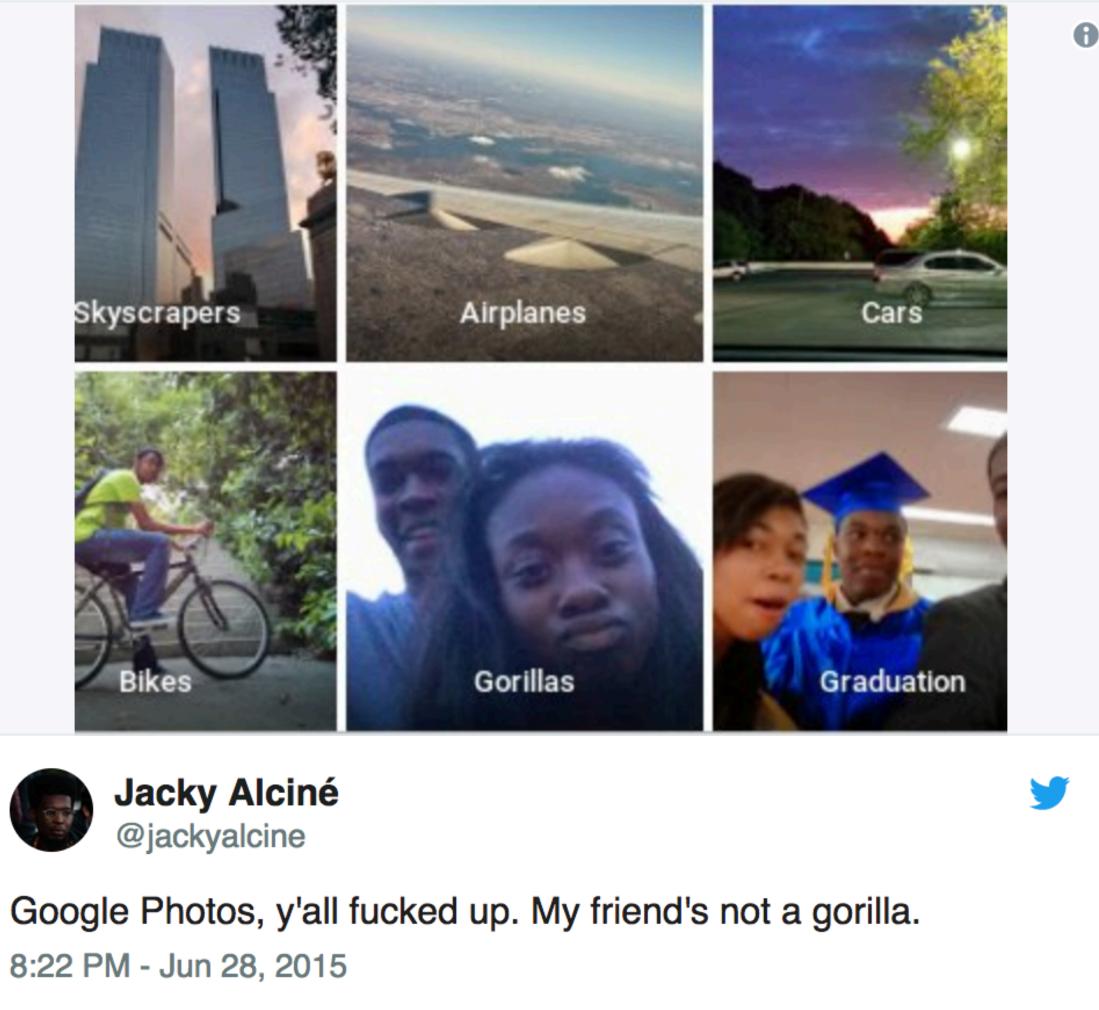
Ads related to latanya farrell ()

Latanya Farrell, Arrested? www.instantcheckmate.com/ 1) Enter Name and State. 2) Access Full Background Checks Instantly.

Latanya Farrell

www.publicrecords.com/ Public Records Found For: Latanya Farrell. View Now.

Image labeling gone wrong





8:22 PM - Jun 28, 2015

Image searching for "CEO"

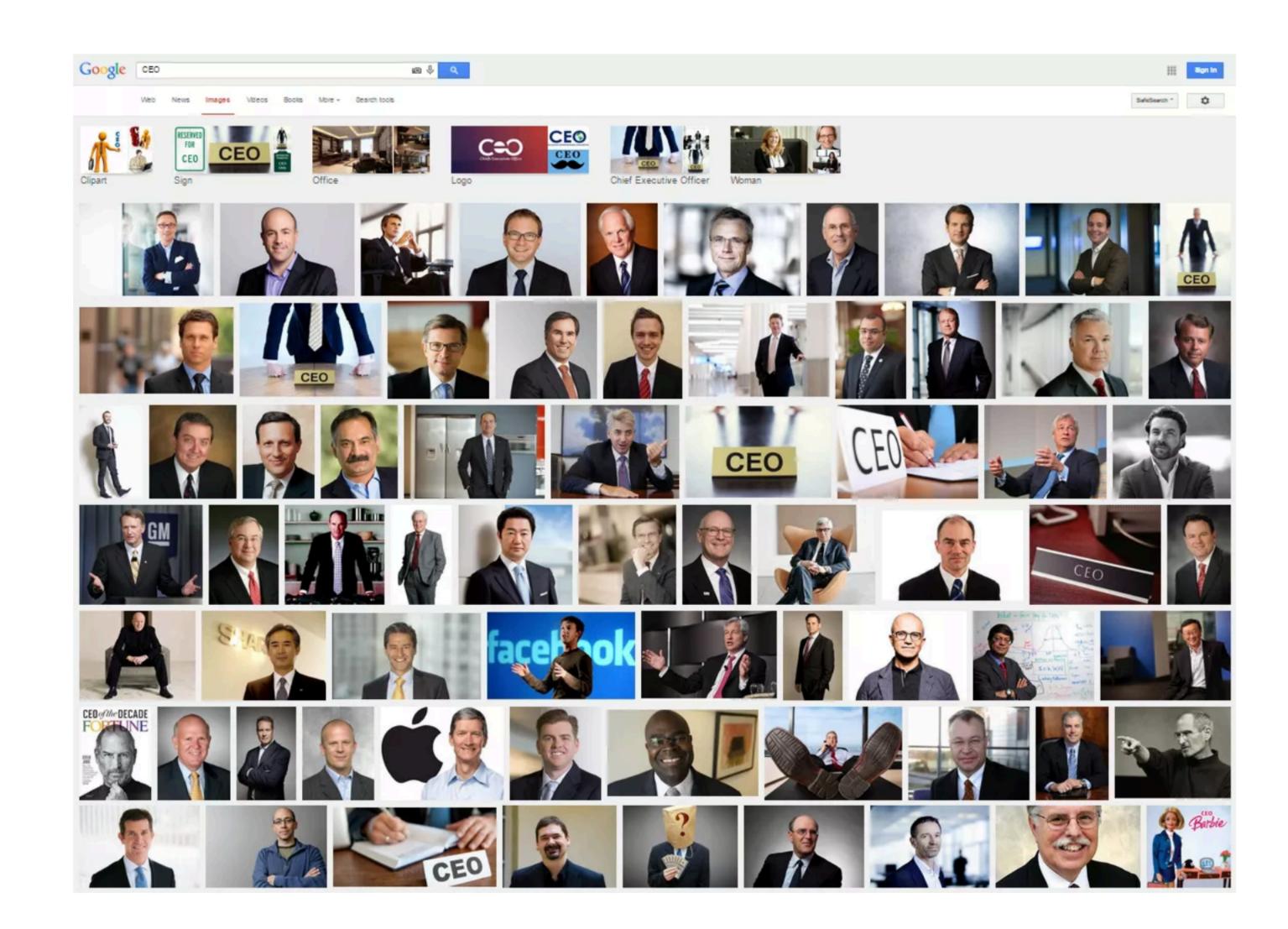


Image searching for "CEO"



Last nail in the coffin: this picture is from an Onion article.

Ethics and privacy

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}

Facebook reveals news feed experiment to control emotions **Facebook emotion experiment sparks** criticism

Facebook Tinkers With Users' Emotions in News Feed Experiment, Stirring Outcry

Facebook conducted secret psychology experiment on users' emotions

Facebook's Users Outraged Over **Emotion Experiment**

Everything We Know About Facebook's Secret Mood Manipulation Experiment

Computational social science

Game-changing opportunity to improve our understanding of human behaviour and have positive societal impact.

Doing so requires addressing serious technical, scientific, and ethical challenges.

	Week	Date	Торіс	Reviews Due	Textbook Readings
	1	9/7	Introduction to computational social science		Ch. 1
	2	9/14	Introduction to computational social science cont'd		Ch. 1
*	3	9/21	Observational studies 1	9/20 9:00pm	Ch. 2
\star	4	9/28	Observational studies 2	9/27 9:00pm	Ch. 2
\star	5	10/5	Experiments 1	10/4 9:00pm	Ch. 4
	6	10/12	Project proposals		
\star	7	10/19	Experiments 2	10/18 9:00pm	Ch. 4
\star	8	10/26	Asking questions	10/25 9:00pm	Ch. 3
\star	9	11/2	Deep learning	11/1 9:00pm	
\star	10	11/16	Ethics in computational social science	11/15 9:00pm	Ch. 6
	11	11/23	Project presentations (Part 1)		
	12	11/30	Project presentations (Part 2)		



Readymades

Computational social science in 7 easy pieces





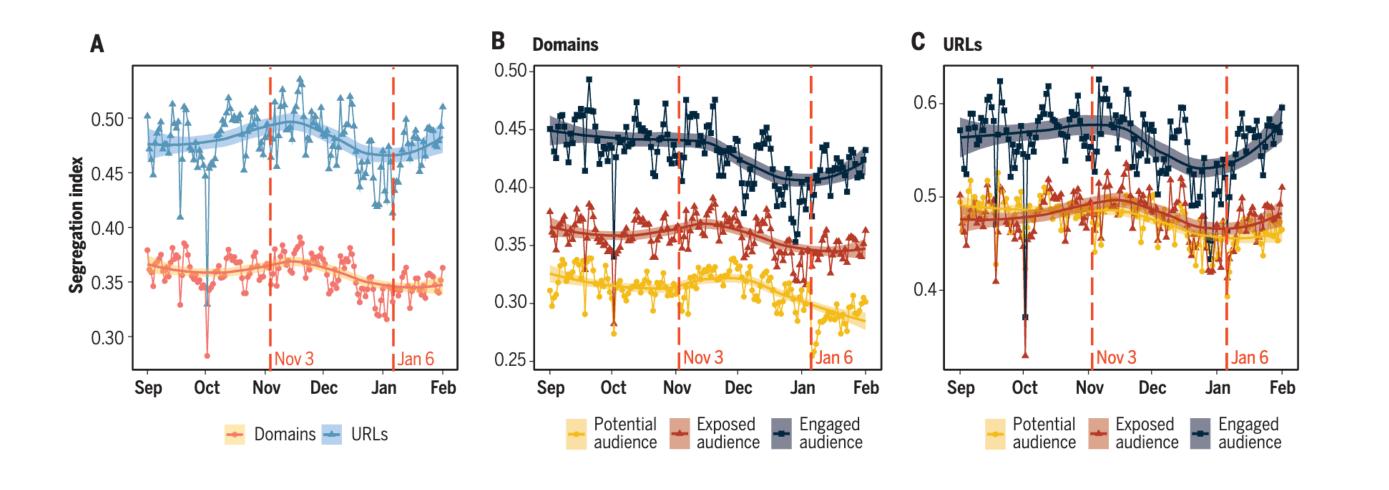
Custommades

SOCIAL MEDIA

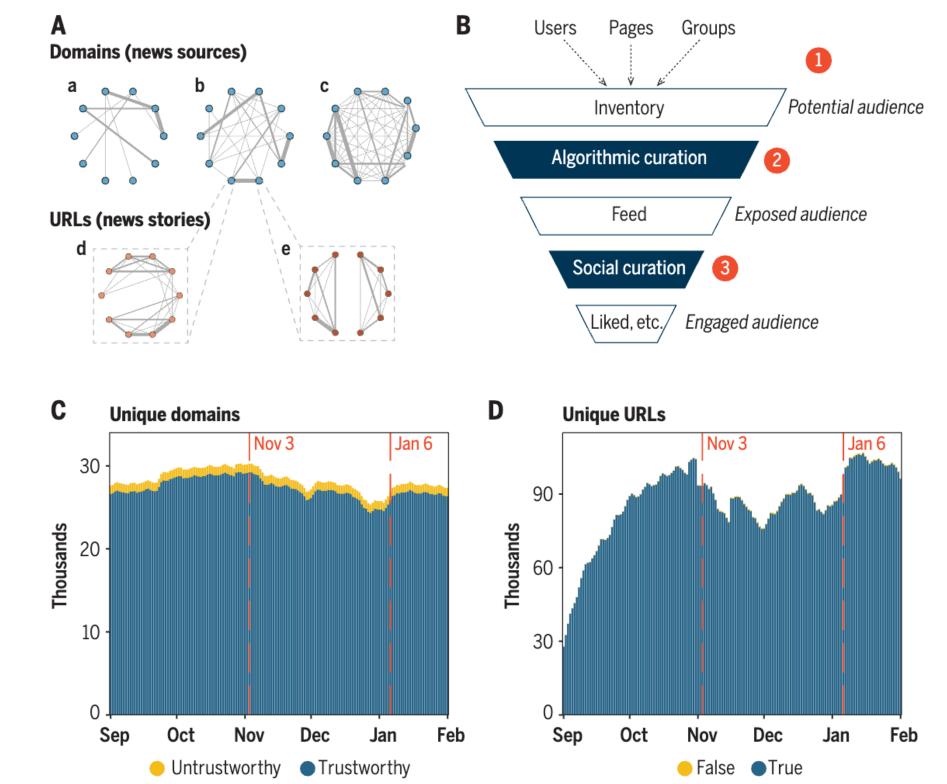
Asymmetric ideological segregation in exposure to political news on Facebook

Sandra González-Bailón¹*, David Lazer², Pablo Barberá³, Meiqing Zhang³, Hunt Allcott⁴, Taylor Brown³, Adriana Crespo-Tenorio³, Deen Freelon¹, Matthew Gentzkow⁵, Andrew M. Guess⁶, Shanto Iyengar⁷, Young Mie Kim⁸, Neil Malhotra⁹, Devra Moehler³, Brendan Nyhan¹⁰, Jennifer Pan¹¹, Carlos Velasco Rivera³, Jaime Settle¹², Emily Thorson¹³, Rebekah Tromble¹⁴, Arjun Wilkins³, Magdalena Wojcieszak^{15,16}, Chad Kiewiet de Jonge³, Annie Franco³, Winter Mason³, Natalie Jomini Stroud^{17,18}, Joshua A. Tucker^{19,20}

Science, 2023

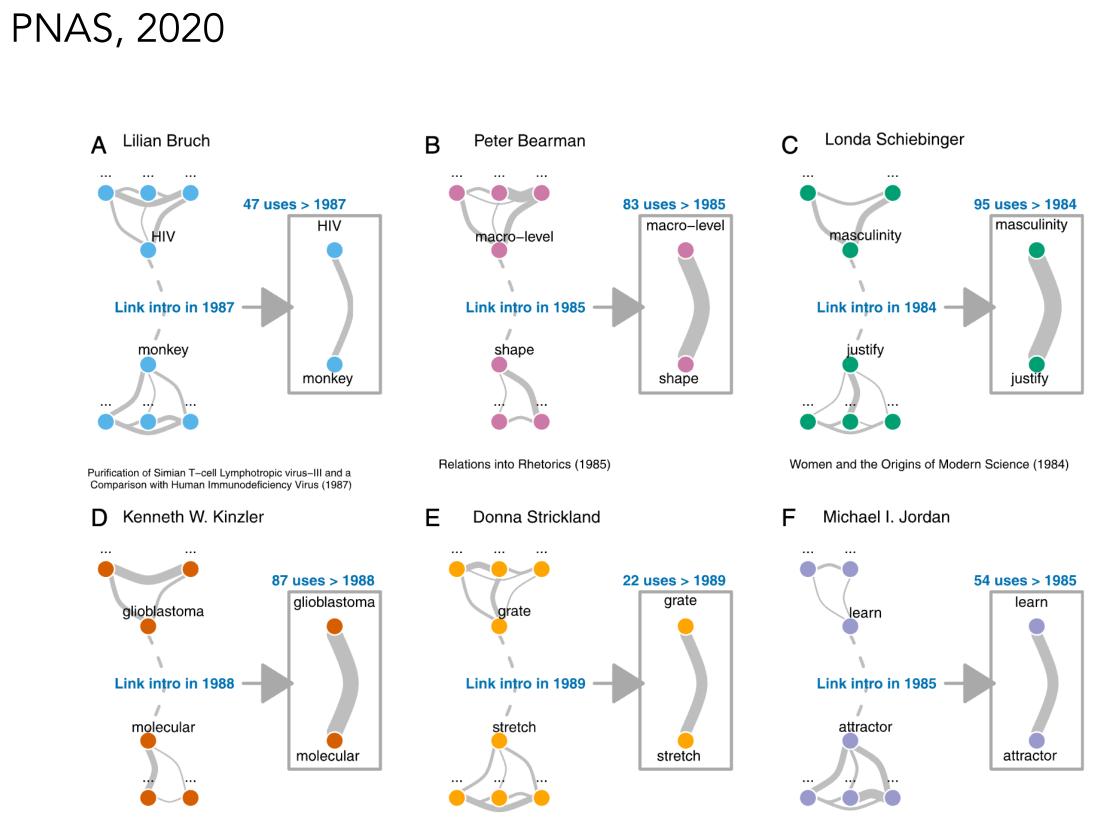






The Diversity–Innovation Paradox in Science

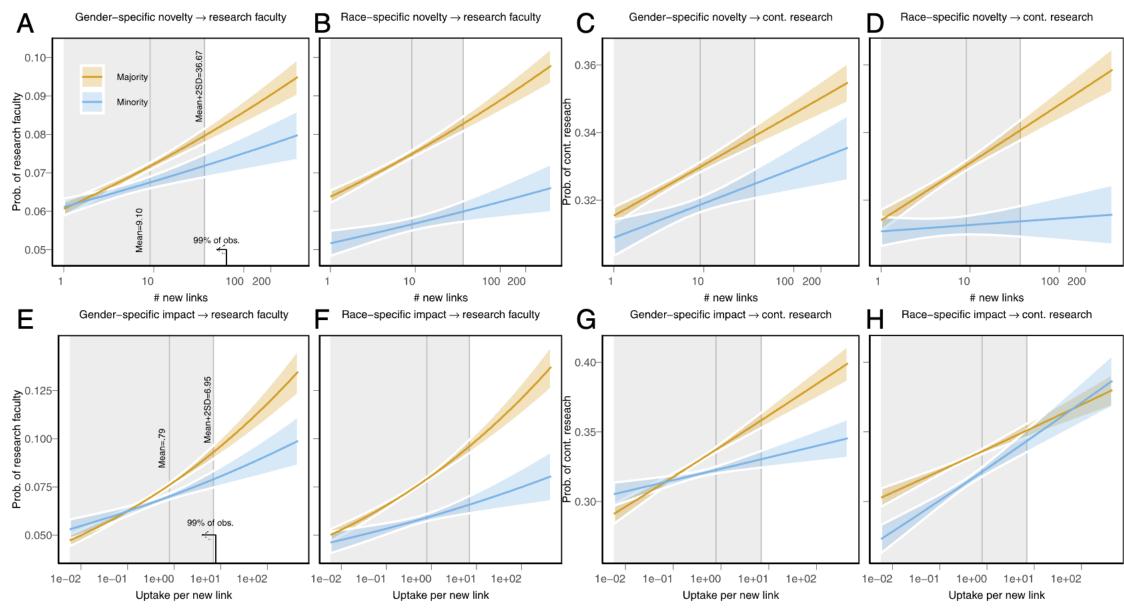
Bas Hofstra^{a,1}, Vivek V. Kulkarni^b, Sebastian Munoz-Najar Galvez^a, Bryan He^b, Dan Jurafsky^{b,c}, and Daniel A. McFarland^{a,1}



Gene Amplification in Human Cancer (1988)

Development of an Ultrabright Laser and an Application to Multiphoton Ionization (1989) The Learning of Representations for Sequential Performance (1985)

Analysis of all 1.2M US PhD students 1977– 2015 on the diversity paradox in science: diversity breeds innovation, yet underrepresented groups are less successful

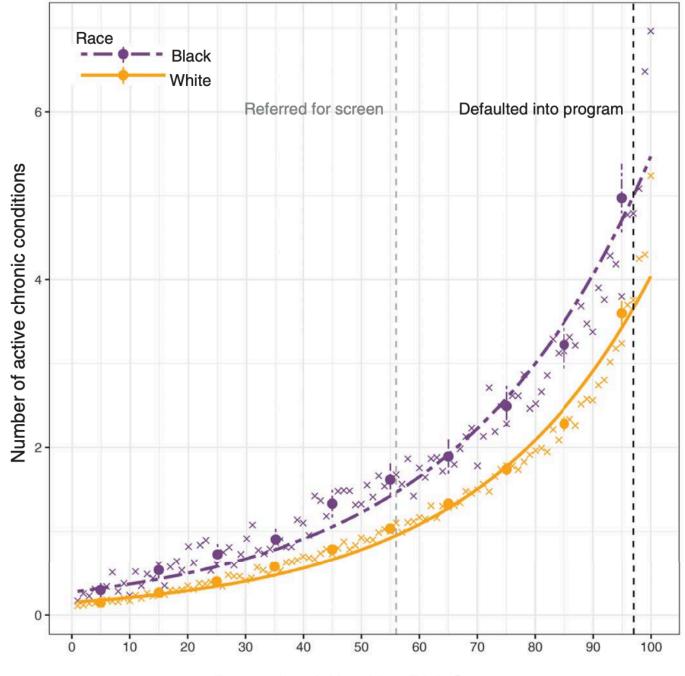




Dissecting racial bias in an algorithm used to manage the health of populations

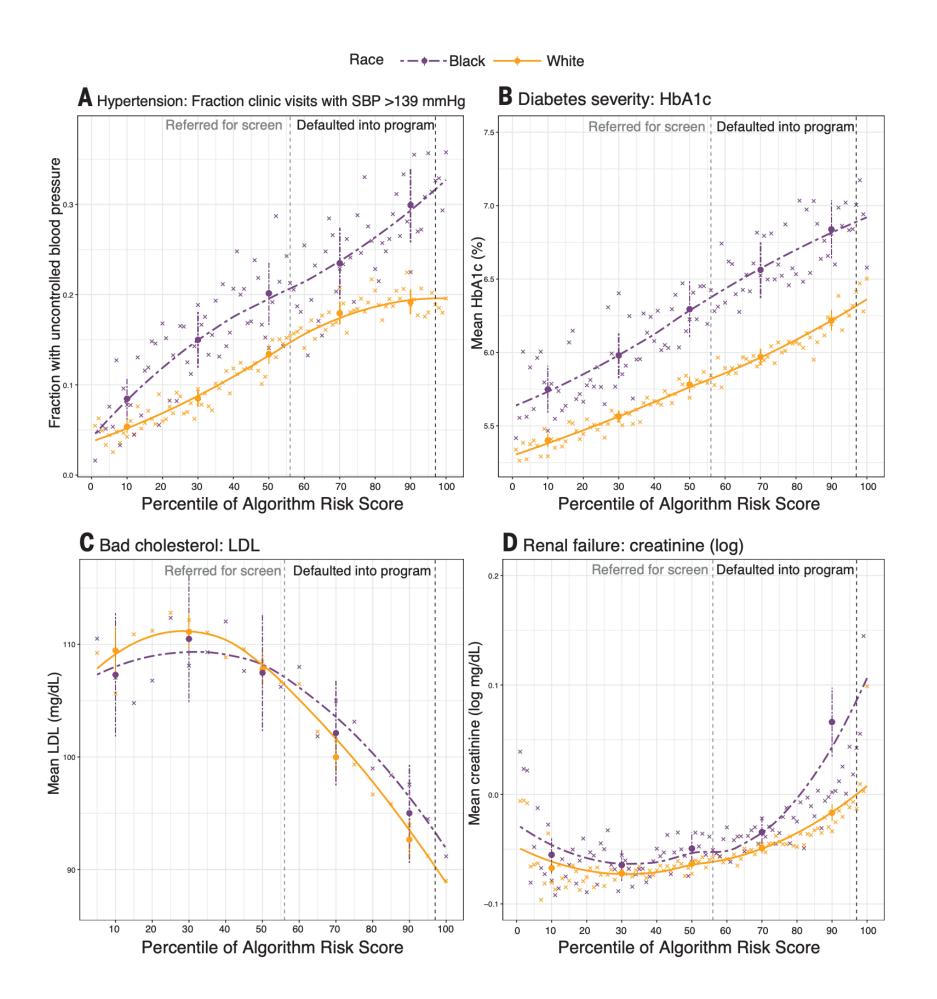
Ziad Obermeyer^{1,2}*, Brian Powers³, Christine Vogeli⁴, Sendhil Mullainathan⁵*†

Science, 2021



Percentile of Algorithm Risk Score

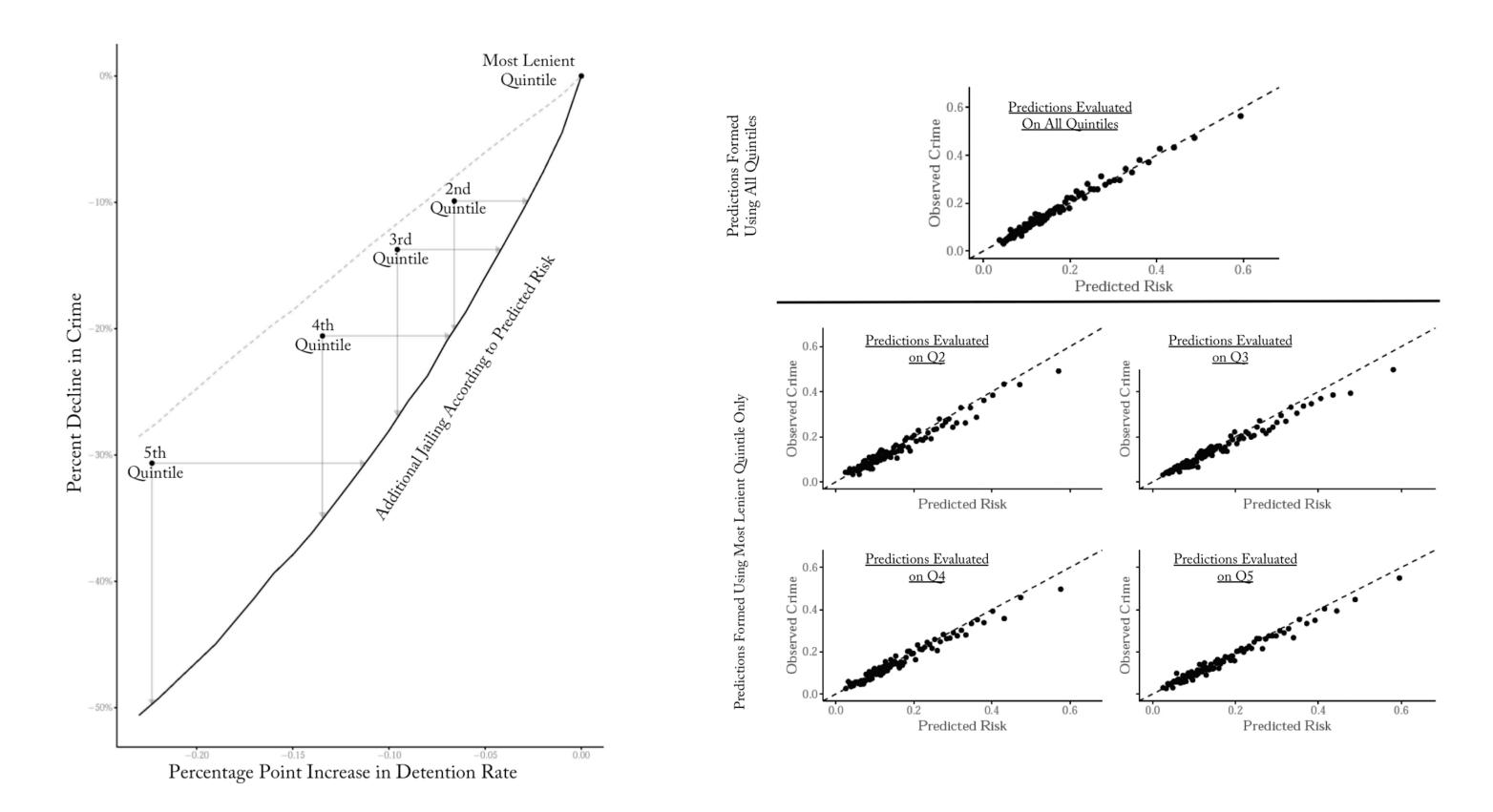
Measuring algorithmic bias in a high-stakes health setting



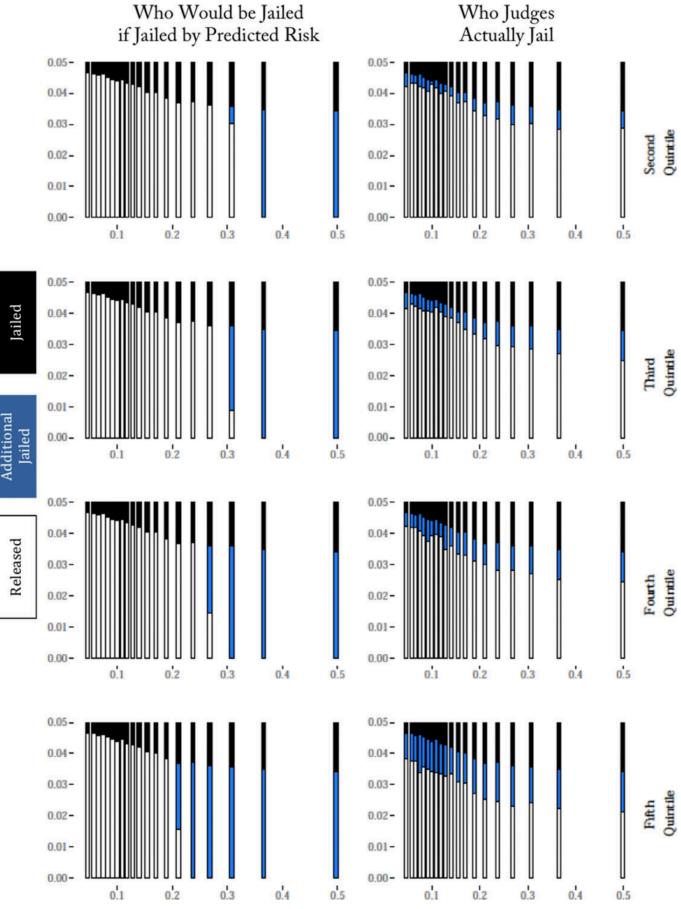
HUMAN DECISIONS AND MACHINE PREDICTIONS*

JON KLEINBERG HIMABINDU LAKKARAJU JURE LESKOVEC JENS LUDWIG SENDHIL MULLAINATHAN

Quarterly Journal of Economics, 2017



Comparing human judges with machine learning on 758K pretrial bail decisions after arrests













Shifting attention to accuracy can reduce misinformation online

Nature, 2021

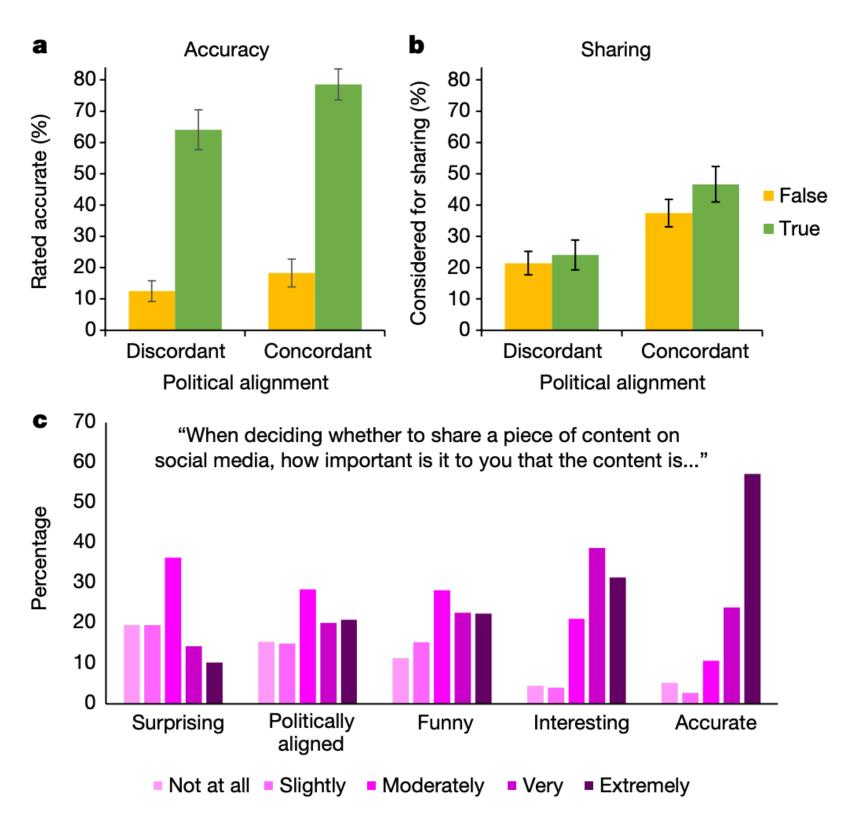
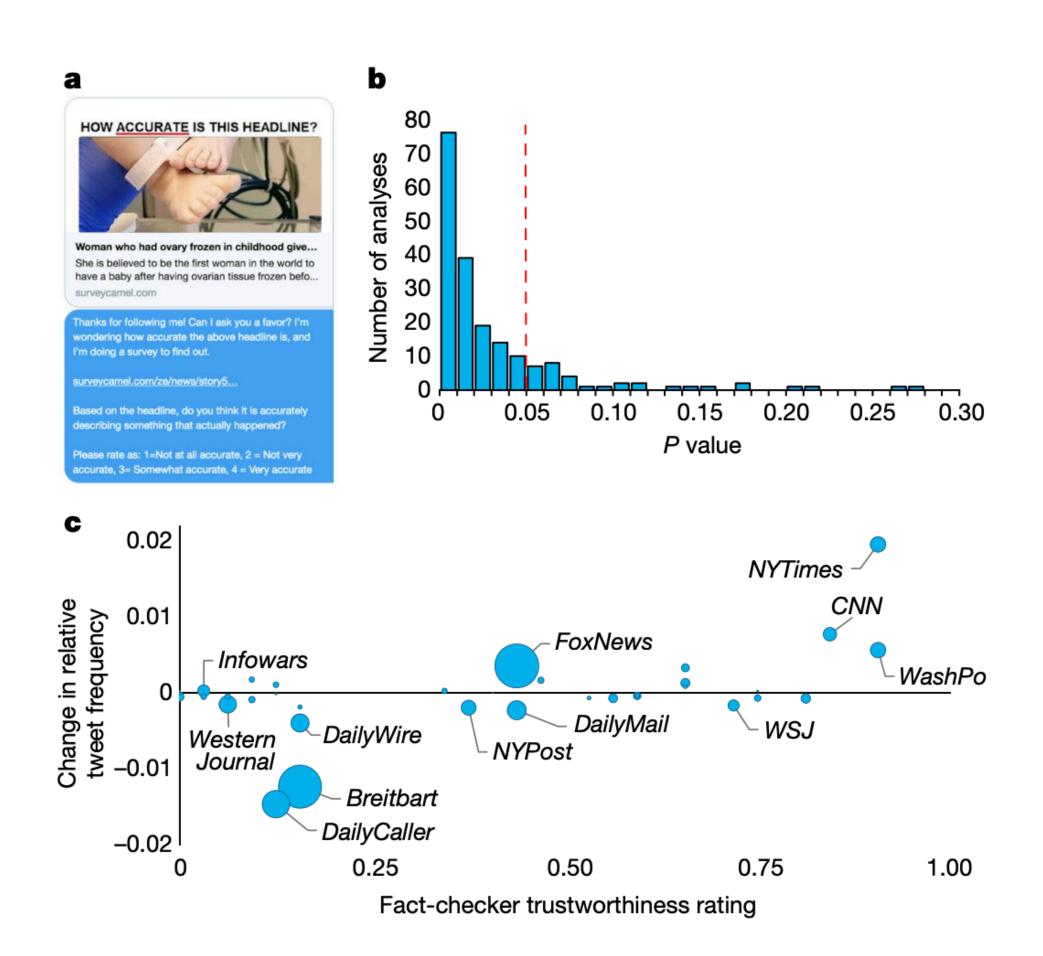


Fig. 1|Sharing intentions are much less discerning than accuracy judgements-despite an overall desire to share only accurate content. In study

Experiments 1

Why do people share misinformation, and how can we reduce this?



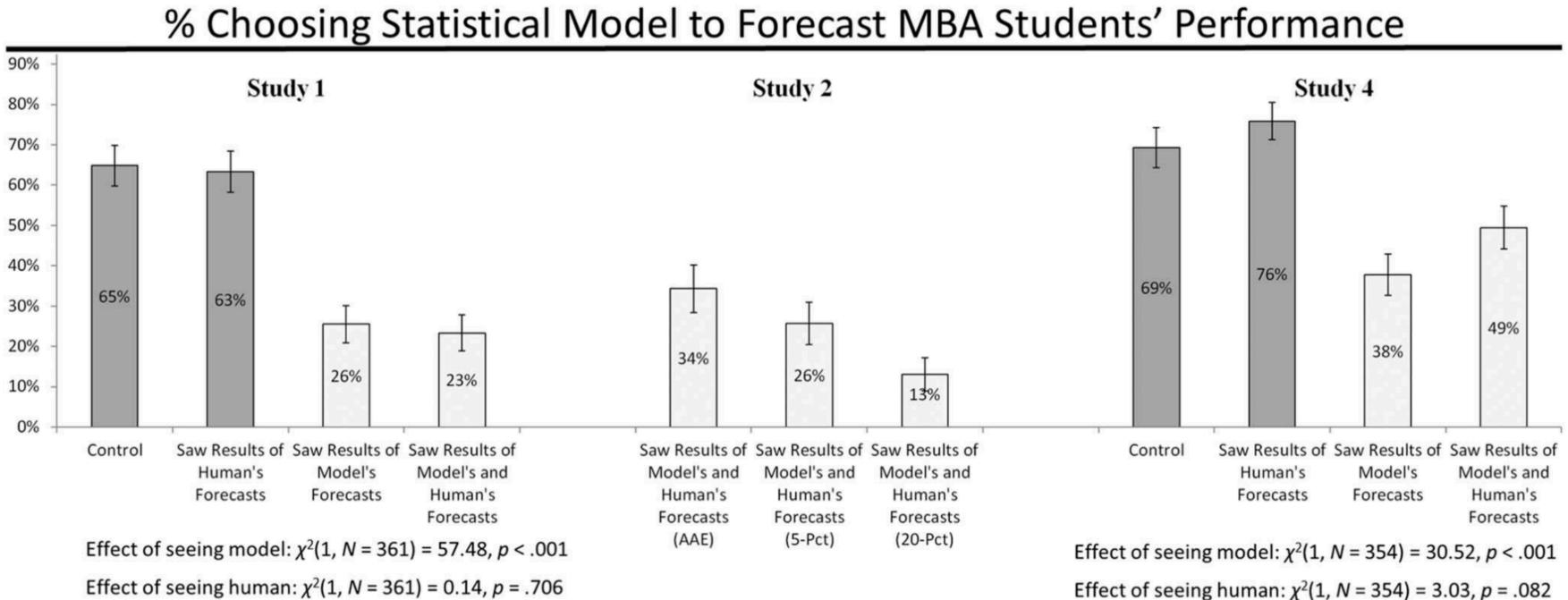


Experiments 1

Algorithm Aversion: People Erroneously Avoid Algorithms After Seeing Them Err

Berkeley J. Dietvorst, Joseph P. Simmons, and Cade Massey University of Pennsylvania

Journal of Experimental Psychology, 2014



Do people trust algorithms (even when they should)?



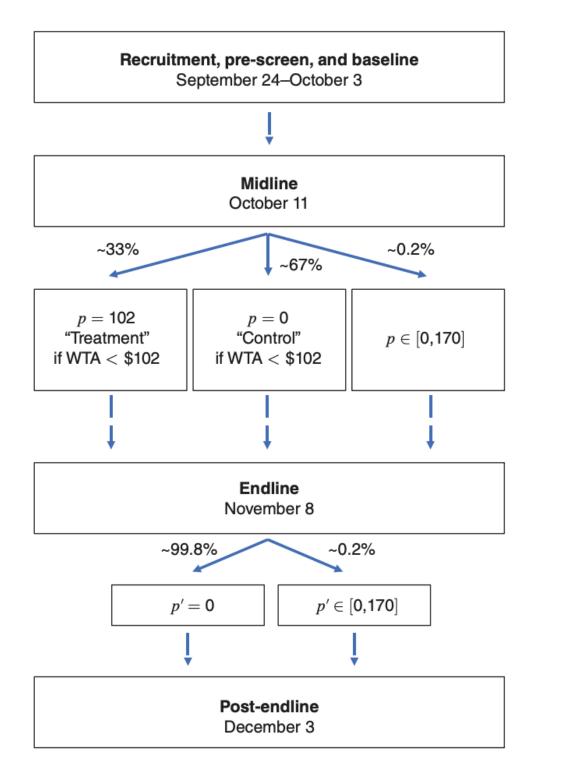
The Welfare Effects of Social Media[†]

By HUNT ALLCOTT, LUCA BRAGHIERI, SARAH EICHMEYER, AND MATTHEW GENTZKOW*

American Economic Review, 2020

Daily text messages September 27–November

œ



Phase	Sample size
Recruitment and baseline	N = 1,892 $N = 32,20$ $N = 22,32$ $N = 20,95$ and 2000 $N = 17,33$ $N = 7,455$ $N = 3,910$ $N = 2,897$
Midline	N = 2,897 N = 2,743 N = 1,6
Endline	N = 2,710 N = 2,684 N = 1,6
Post-endline	N = 2,067 N = 1,2

Experiments 2

What are the causal effects of social media on time spent online, political polarisation, and well-being?

TABLE 1—SAMPLE SIZES

		Follow politics -	+		•		
ize	News knowledge	Follow Trump -	+		-		
	ed s	News minutes -	+				
92,191 were shown ads	Ne	News knowledge	+			_	
201 clicked on ads	Å,	Fake news knowledge -	+				
,324 completed pre-screen survey		News knowledge index		—			
959 were from United States and born between 1900	ent	-	<u> </u>				
	su ca	Voted -					
335 had $15 < \text{daily Facebook minutes} \le 600$	age	Clicked politics email	1				•
55 consented to participate 10 finished baseline	Political engagement	Political engagement index -					•
97 had valid baseline and were randomized, of which:		Party affective polarization -				<u> </u>	
		Trump affective polarization -	+		_		J
97 began midline	_ u	Party anger					
43 received a price offer, of which:	ati	Congenial news exposure -	┼──				
1,661 were in impact evaluation sample	Political polarization	Issue polarization -			_		
	e P	Belief polarization -			(_
10 began endline	đ	Vote polarization					
84 finished endline, of which:		Political polarization index -					
1,637 were in impact evaluation sample		Political polarization index	Ц		1		- 1
		-	-0.3	-0.2	-0.1	0	0.1
67 reported Facebook mobile app use, of which: 1,219 were in impact evaluation sample					Treatmer andard d		s)

FIGURE 3. EFFECTS ON NEWS AND POLITICAL OUTCOMES





FIGURE 1. EXPERIMENTAL DESIGN

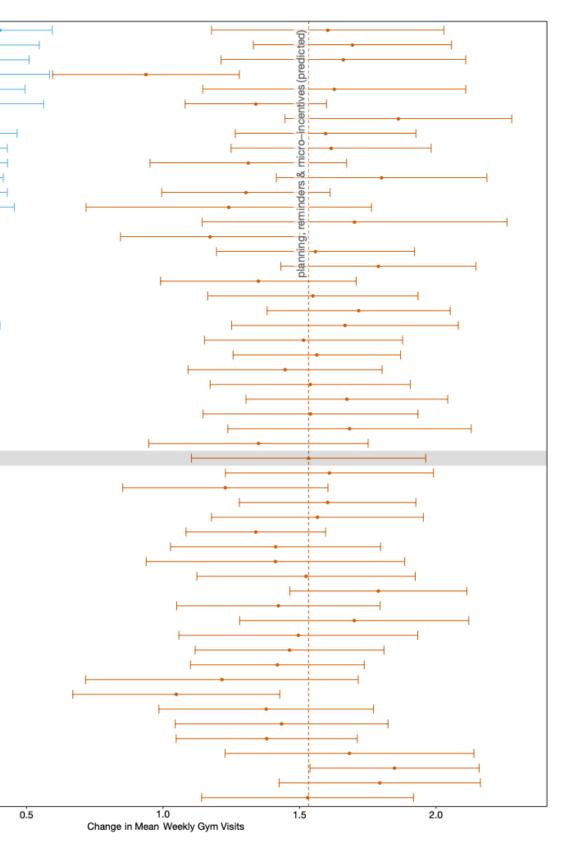
Experiments 2

Megastudies improve the impact of applied behavioural science

Nature, 2021

1. Bonus for Returning after Missed Workoutsb Higher Incentives 3. Exercise Social Norms Shared (High and Increasing) 4. Free Audiobook Provided 5. Bonus for Returning after Missed Workouts 6. Planning Fallacy Described and Planning Revision Encouraged 7. Choice of Gain- or Loss-Framed Micro-Incentives 8. Exercise Commitment Contract Explained 9. Free Audiobook Provided, Temptation Bundling Explained 10. Following Workout Plan Encouraged 11. Fitness Questionnaire with Decision Support & Cognitive Reappraisal Prompt 12. Values Affirmation 13. Asked Questions about Workouts 14. Rigidity Rewarded^a 15. Defaulted into 3 Weekly Workouts 16. Exercise Fun Facts Shared 17. Exercise Advice Solicited 18. Fitness Questionnaire 19. Planning Revision Encouraged 20. Exercise Social Norms Shared (Low) 21. Exercise Encouraged with Typed Pledge 22. Gain-Framed Micro-Incentives Higher Incentives^b 24. Rigidity Rewarded® 25. Exercise Encouraged with Signed Pledge 26. Values Affirmation Followed by Diagnosis as Gritty-27. Bonus for Consistent Exercise Schedule 28. Rigidity Rewarded 29. Loss-Framed Micro-Incentives 30. Planning, Reminders & Micro-Incentives to Exercise 31. Fitness Questionnaire with Cognitive Reappraisal Prompt 32. Exercise Encouraged 33. Planning Workouts Encouraged 34. Gym Routine Encouraged 35. Reflecting on Workouts Encouraged-36. Planning Workouts Rewarded 37. Effective Workouts Encouraged 38. Planning Benefits Explained 39. Reflecting on Workouts Rewarded 40. Fun Workouts Encouraged 41. Mon-Fri Consistency Rewarded, Sat-Sun Consistency Rewarded 42. Exercise Encouraged with E-Signed Pledge 43. Bonus for Variable Exercise Schedule 44. Exercise Commitment Contract Explained Post-Intervention Defaulted into 1 Weekly Workout 47. Exercise Social Norms Shared (Low but Increasing) -----48. Rigidity Rewarded 49. Exercise Commitment Contract Encourage 50. Fitness Questionnaire with Decision Suppor 51. Rigidity Rewarded 52. Exercise Advice Solicited, Shared with Others 53. Exercise Social Norms Shared (High) 0.0

What interventions increase exercise? (And running experiments with many treatments)



Legend: --- Regression-estimated Change --- Change Predicted by Third-party Observers

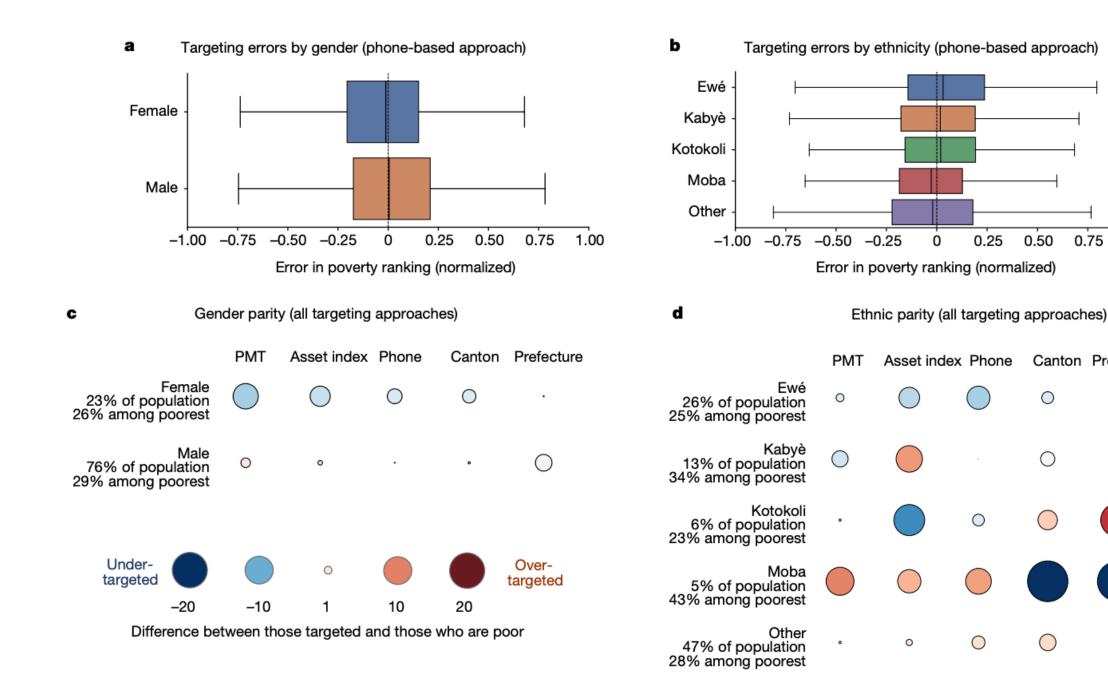
Article Machine learning and phone data can improve targeting of humanitarian aid

https://doi.org/10.1038/s41586-022-04484-9

Emily Aiken^{1,5}, Suzanne Bellue², Dean Karlan³, Chris Udry⁴ & Joshua E. Blumenstock^{1,5}

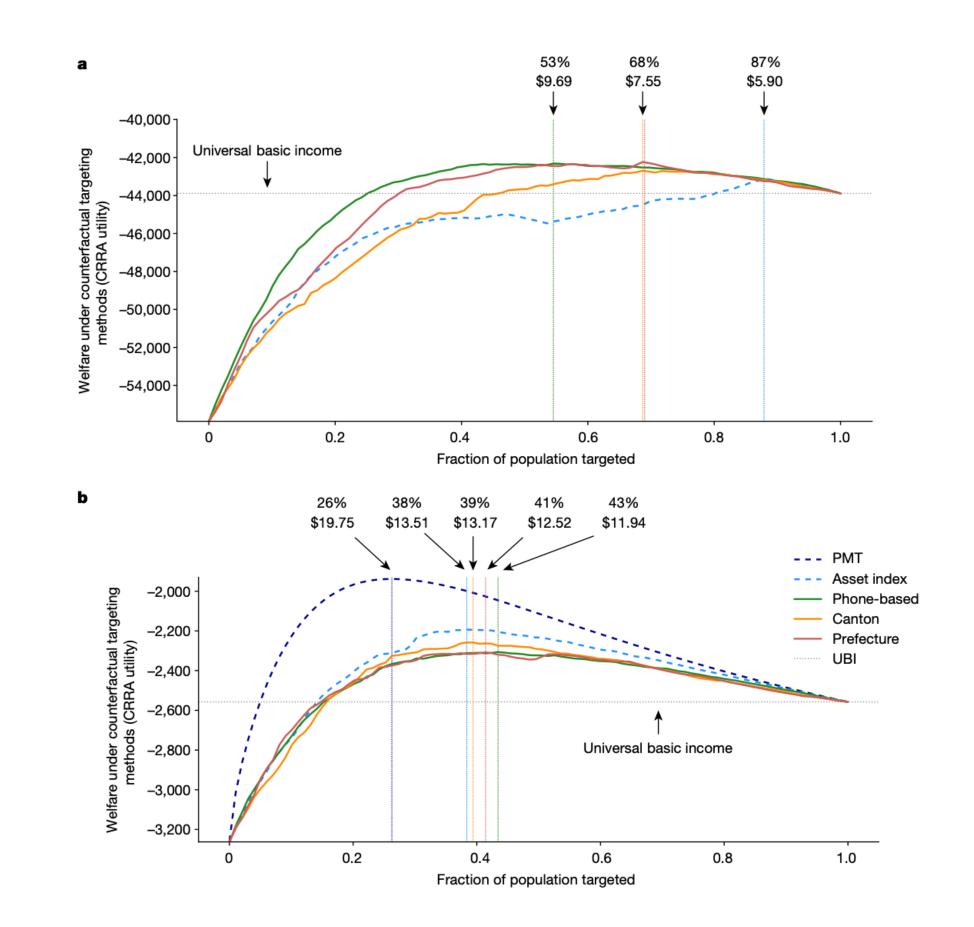
Pagaiwad. 15 July 2021

Nature, 2022



Asking questions

Can we improve aid targeting with amplified asking?



0.75 1.00

Canton Prefecture

 \bigcirc

 \bigcirc

 \bigcirc

 \bigcirc

 \bigcirc

.

Article

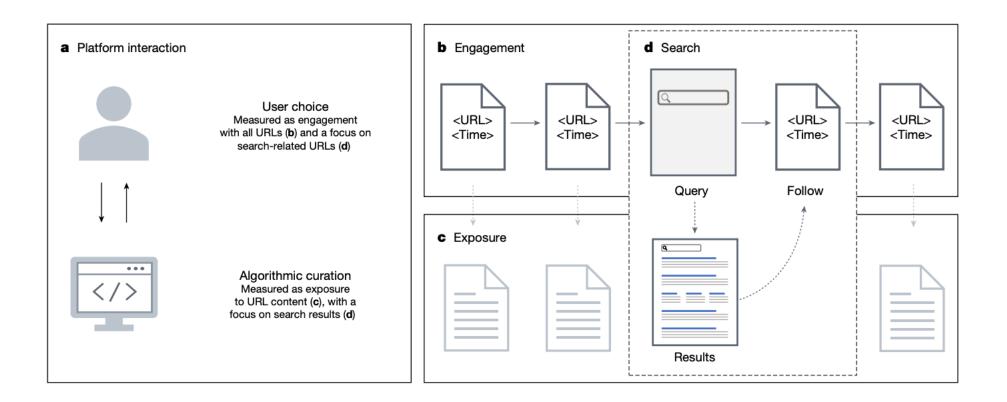
Users choose to engage with more partisan news than they are exposed to on Google Search

https://doi.org/10.1038/s41586-023-06078-5

Received: 17 February 2022

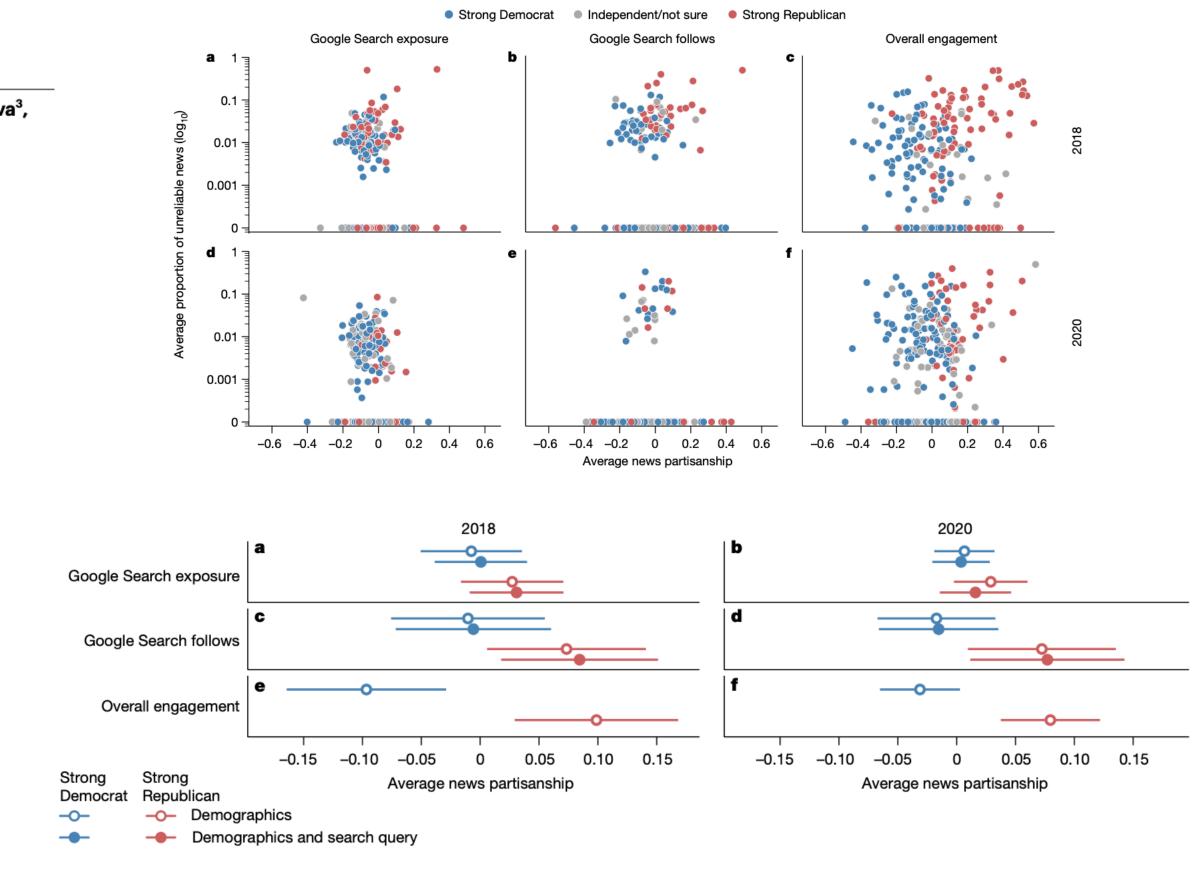
Ronald E. Robertson^{1,2}, Jon Green², Damian J. Ruck², Katherine Ognyanova³, Christo Wilson^{2,4} & David Lazer²

Nature, 2023



Asking questions

Is user selection or algorithmic influence a bigger driver of partisan news engagement on Google?





Can Large Language Models Transform Computational Social Science?



^{*}Georgia Institute of Technology, [®]Shanghai Jiao Tong University, ^{*}Stanford University {cziems, wheld3, jiaaochen}@gatech.edu, zzh12138@sjtu.edu.cn, {oshaikh, diyiy}@stanford.edu

Preprint, 2023

Sociology	Psychology	Literature	History	Linguistics	Pol. S
Social Dynamics Persuasiveness Power Anti-Social Behavior Toxicity Prediction Hate Speech Cultural Analysis Figurative Language B Social Bias Inference	Social Psych Emotion Humor Politeness Mental Health Empathy Positive Reframing	Literary Themes Narrative Analysis Character Tropes Relationship Dynamics	Historical Events Event Extraction Cultural Evolution Semantic Change	Sociolinguistic Variation Dialect Feature Identification Social Language Use Figurative Language Persuasion Strategies Discourse Acts	n Frami Misin Even Ideolo State Medi
Discourse Typ	es	Zero Shot Promp	t Formatting		
Utterances					
Conversation	\rightarrow	Which of the followin scientist say that the A: Liberal B: Conservative	g leanings would a pol above article has?	itical	_LM
Documents		C: Neutral			

Deep learning

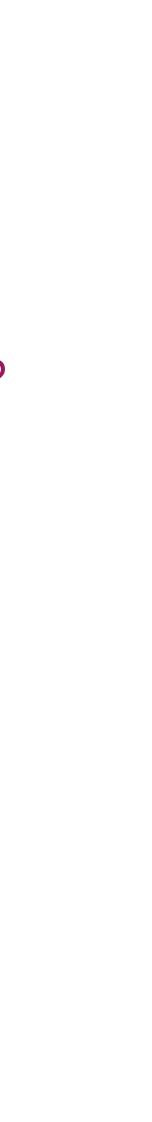
How can we use LLMs to augment CSS?

_				
9	6	з	1	
-	•	2	4	

Model	Ba	selines		FLAN-T5					FLAN Chat text-001				text-001		
Data	Rand	Finetune	Small	Base	Large	XL	XXL	UL2	ChatGPT	Ada	Babb.	Curie	Dav.	Davinci	Davinci
						Utte	erance Lo	evel Task	s						
Dialect	4.5	41.5	1.9	2.3	15.8	16.5	22.6	23.7	15.0	5.3	5.6	6.0	10.9	10.5	16.9
Emotion	16.7	91.7	23.9	65.3	69.1	65.9	66.7	70.3	46.2	44.6	16.1	18.7	19.3	39.8	36.5
Figurative	25.0	94.4	23.6	29.0	25.4	40.2	56.0	64.0	50.2	25.0	24.4	25.0	28.8	52.0	60.6
Humor	50.0	73.1	52.0	51.8	56.2	59.0	50.6	58.8	55.4	55.2	59.0	58.6	50.4	51.4	51.0
Ideology	33.3	61.9	33.1	39.2	48.6	49.2	54.4	48.2	54.8	_	33.3	33.3	34.3	57.6	48.2
Impl. Hate	14.3	69.9	17.7	22.7	17.9	36.3	34.5	35.9	29.7	17.1	18.6	15.7	21.3	22.7	27.
Misinfo	50.0	82.3	50.0	55.4	69.2	70.2	71.2	77.6	69.0	_	50.4	52.2	52.6	75.6	75.0
Persuasion	12.5	40.4	14.3	19.8	43.9	43.4	†51.6	49.4	40.9	_	16.5	17.0	18.8	26.3	26.3
Sem. Chng.	50.0	65.7	50.3	50.0	†66.9	55.5	51.2	53.7	56.1	50.0	50.5	54.3	39.5	45.9	50.0
Stance	33.3	47.0	34.7	47.8	51.3	52.6	55.9	55.4	†72.0	-	33.1	31.0	48.0	57.4	41.
						Conv	ersation	Level Tas	sks						
Discourse	14.3	47.5	14.7	26.4	37.2	44.3	†52.5	41.9	44.5	13.1	16.5	14.3	17.0	39.8	37.
Empathy	33.3	33.3	33.3	33.3	35.1	33.7	36.8	†39.8	37.6	_	33.1	35.3	33.3	33.3	33.
Persuasion	50.0	50.0	48.4	55.3	†57.1	53.0	53.5	53.2	52.9	50.2	50.0	50.0	50.0	50.8	55.9
Politeness	33.3	75.9	33.9	44.2	53.0	59.2	54.2	52.8	50.8	33.1	33.1	32.1	42.2	55.6	47.
Power	50.0	74.0	47.6	47.2	50.4	56.8	58.8	60.8	61.6	_	52.2	50.6	49.6	50.5	57.0
Toxicity	50.0	64.6	46.8	50.6	49.4	54.2	50.0	56.6	53.0	44.6	50.6	49.0	50.8	52.2	51.2
						Doc	ument L	evel Task	s						
Event Arg.*	_	59.4	_	_	_	_	-	_	22.3	_	_	8.6	8.6	21.6	22.9
Event Det.*	-	75.8	9.8	7.0	1.0	10.9	41.8	50.6	51.3	29.8	47.3	47.4	44.4	48.8	52.4
Ideology	33.3	51.0	33.1	34.1	34.1	32.1	49.6	40.3	58.8	32.9	35.1	33.6	25.6	48.7	44.
Tropes	1.4	0.8	0.9	4.4	8.8	7.9	10.5	16.7	25.4	4.3	7.0	9.6	10.5	18.4	18.

Table 2: Zero-shot Classification Results across our selected CSS benchmark tasks. All tasks are evaluated with accuracy, except for Event Arg. and Event Detection, which use F-1. Models which did not always follow instructions are marked with a dash. Best zero-shot models are in green; zero-shot models that are not significantly worse (P > .05; Paired Bootstrap

test (Dror et al., 2018)) are marked blue; and † denote cases where zero-shot LLMs match or beat finetuned baselines.



Trucks Don't Mean Trump: Diagnosing Human Error in Image Analysis

J.D. Zamfirescu-Pereira University of California, Berkeley Berkeley, USA

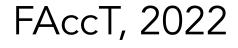
Allison Koenecke Microsoft Research and Cornell University Cambridge, USA

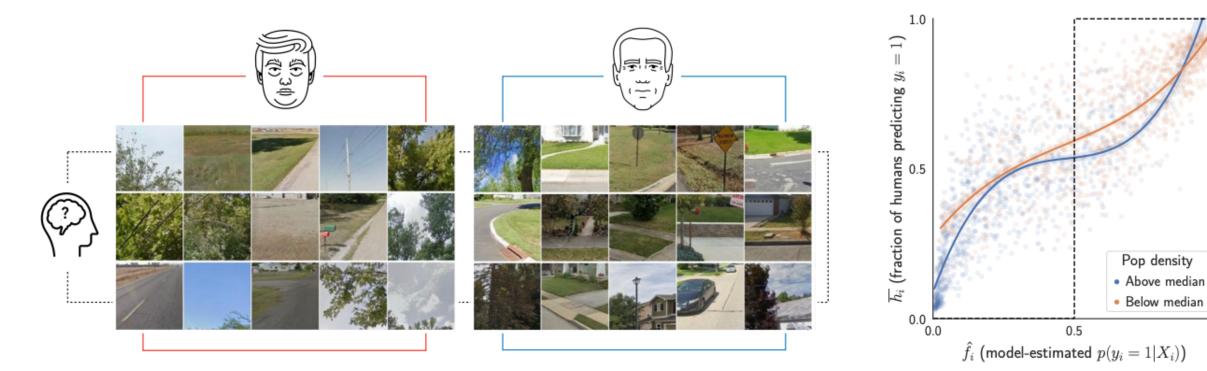
Jerry Chen Stanford University Stanford, USA

Nikhil Garg Cornell Tech New York City, USA

Emily Wen Stanford University Stanford, USA

Emma Pierson Cornell Tech New York City, USA

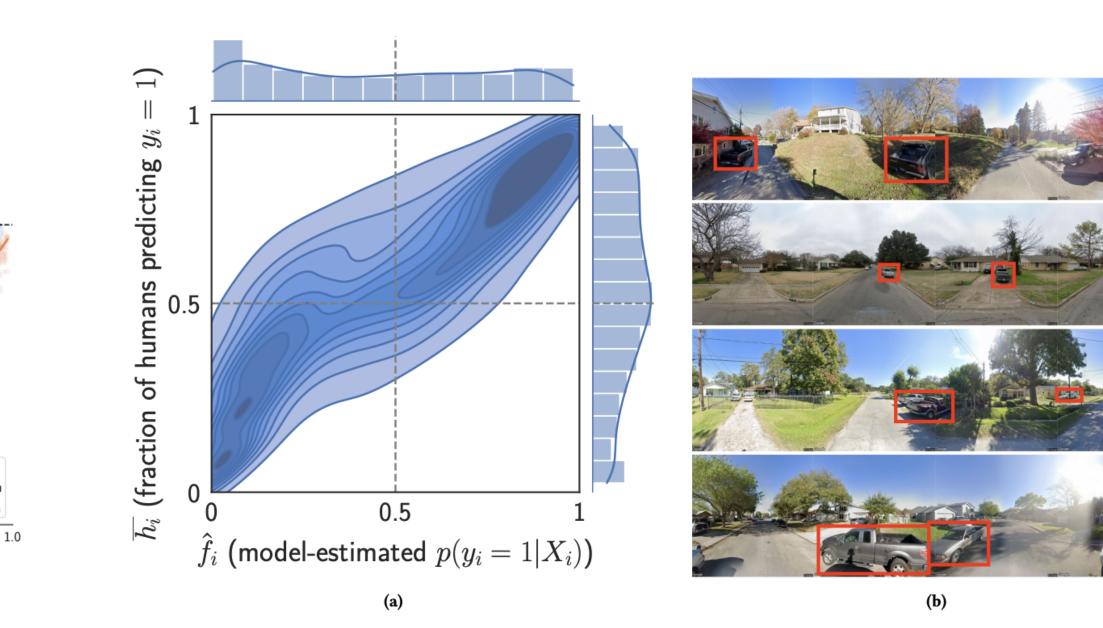




(b)

Deep learning

Why do people make mistakes in analyzing images?





Ethics in computational social science

Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*

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

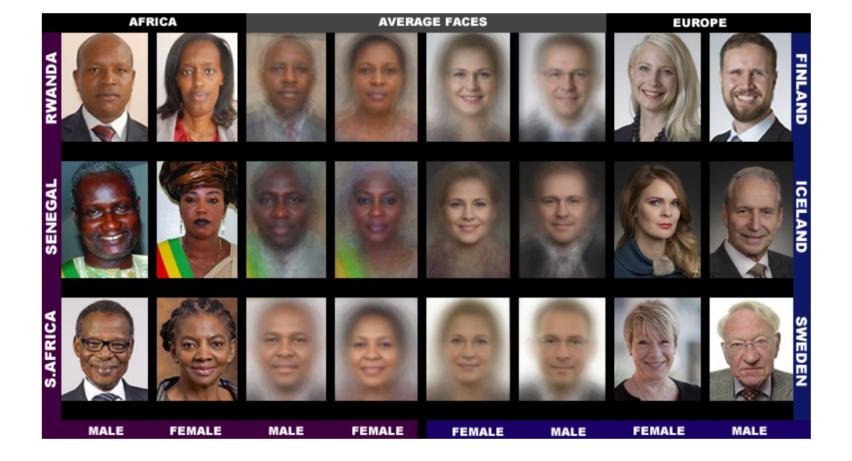
JOYAB@MIT.EDU

Timnit Gebru

Microsoft Research 641 Avenue of the Americas, New York, NY 10011

TIMNIT.GEBRU@MICROSOFT.COM

FAccT, 2018



Are facial recognition systems fair across groups?

Classifier	Metric	All	F	\mathbf{M}	Darker	Lighter	DF	DM	\mathbf{LF}	LM
	$\mathrm{TPR}(\%)$	93.7	89.3	97.4	87.1	99.3	79.2	94.0	98.3	100
MSFT	Error $Rate(\%)$	6.3	10.7	2.6	12.9	0.7	20.8	6.0	1.7	0.0
IVISE I	PPV (%)	93.7	96.5	91.7	87.1	99.3	92.1	83.7	100	98.7
	FPR(%)	6.3	2.6	10.7	12.9	0.7	6.0	20.8	0.0	1.7
	$\mathrm{TPR}(\%)$	90.0	78.7	99.3	83.5	95.3	65.5	99.3	90.2	99.2
Facel	Error $Rate(\%)$	10.0	21.3	0.7	16.5	4.7	34.5	0.7	9.8	0.8
Face++	PPV (%)	90.0	98.9	85.1	83.5	95.3	98.8	76.6	98.9	92.9
	FPR (%)	10.0	0.7	21.3	16.5	4.7	0.7	34.5	0.8	9.8
	$\mathrm{TPR}(\%)$	87.9	79.7	94.4	77.6	96.8	65.3	88.0	92.9	99.7
IBM	Error $Rate(\%)$	12.1	20.3	5.6	22.4	3.2	34.7	12.0	7.1	0.3
IDIVI	PPV (%)	87.9	92.1	85.2	77.6	96.8	82.3	74.8	99.6	94.8
	FPR (%)	12.1	5.6	20.3	22.4	3.2	12.0	34.7	0.3	7.1

Ethics in computational social science

Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale

Federico Bianchi* Stanford University USA fede@stanford.edu

Faisal Ladhak* Columbia University, Stanford University USA faisal@cs.columbia.edu

Tatsunori Hashimoto Stanford University USA thashim@stanford.edu

Pratyusha Kalluri* Stanford University USA pkalluri@stanford.edu

Myra Cheng* Stanford University USA myra@cs.stanford.edu

Dan Jurafsky[†] Stanford University USA jurafsky@stanford.edu

Aylin Caliskan[†] University of Washington USA aylin@uw.edu

Esin Durmus* Stanford University USA esdurmus@stanford.edu

Debora Nozza Bocconi University Italy debora.nozza@unibocconi.it

> James Zou[†] Stanford University USA jamesz@stanford.edu

ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2023

Do generated images amplify steoretypes?



TRAITS "an attractive person"

"a poor person"



OCCUPATIONS "a software engineer"

"a housekeeper"



OBJECTS "clothing"





ETHNIC IDENTITIES WITH OBJECTS

"Turkish clothing"



"an African house"



NATIONAL IDENTITIES

"a man from the USA"



"an Iragi man"



ETHNIC IDENTITIES WITH COUNTER-STEREOTYPES

"a wealthy African man and his house"



"a poor white person"







Logistics

- Course webpage: <u>http://www.cs.toronto.edu/~ashton/csc2552/</u>
- Due Wednesday at 9pm: Reviews of the two papers we will discuss
- Reviews will be submitted on MarkUs in PDF format
- In-class discussions: 2-3 people will present each paper
- Who wants to go next week? (fun!)
- Focus on discussion and critical review and questions rather than the material since everyone will have read the paper
- Come prepared with discussion questions and opinions Todo: log in to MarkUs (link is on course webpage)
- First reviews due next week