Responsible data science

Because of its tremendous **power**, data science must be used **responsibly**

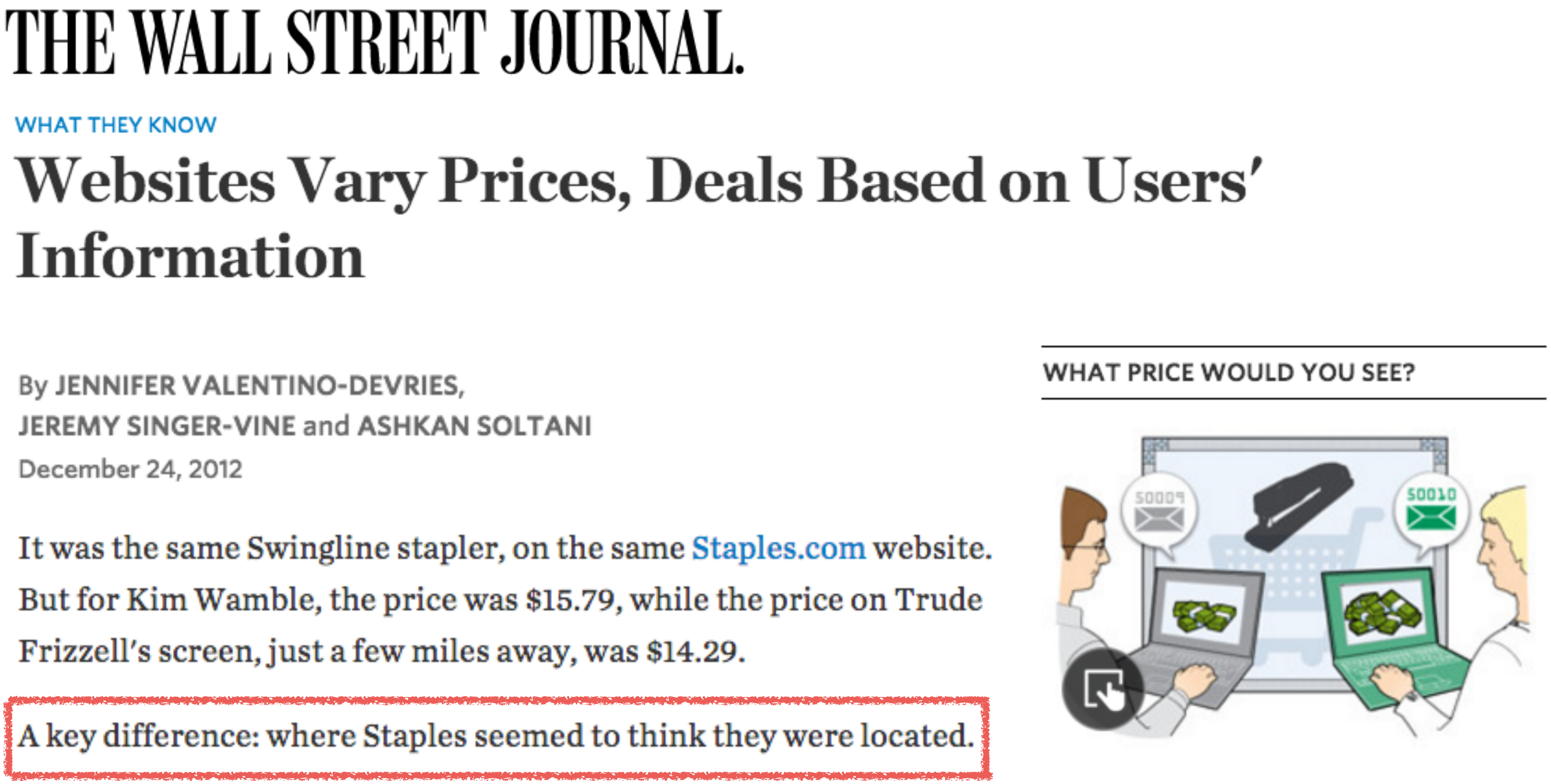
fairness

diversity

transparency

data protection

Online price discrimination



THE WALL STREET JOURNAL.

WHAT THEY KNOW

Websites Vary Prices, Deals Based on Users' Information

By JENNIFER VALENTINO-DEVRIES, JEREMY SINGER-VINE and ASHKAN SOLTANI
December 24, 2012

It was the same Swingline stapler, on the same Staples.com website. But for Kim Wamble, the price was $15.79, while the price on Trude Frizzell's screen, just a few miles away, was $14.29.

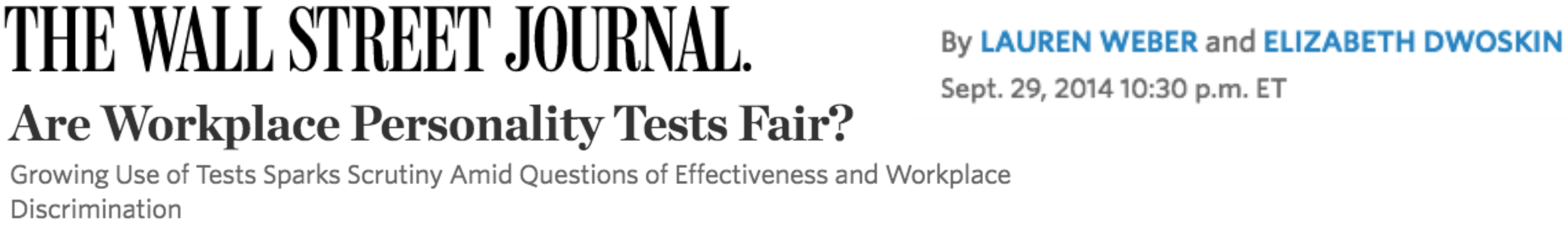
A key difference: where Staples seemed to think they were located.

WHAT PRICE WOULD YOU SEE?

lower prices offered to buyers who live in more affluent neighborhoods

https://www.wsj.com/articles/SB10001424127887323777204578189391813881534

Job-screening personality tests



THE WALL STREET JOURNAL.

By LAUREN WEBER and ELIZABETH DWOSKIN

Sept. 29, 2014 10:30 p.m. ET

Are Workplace Personality Tests Fair?

Growing Use of Tests Sparks Scrutiny Amid Questions of Effectiveness and Workplace Discrimination

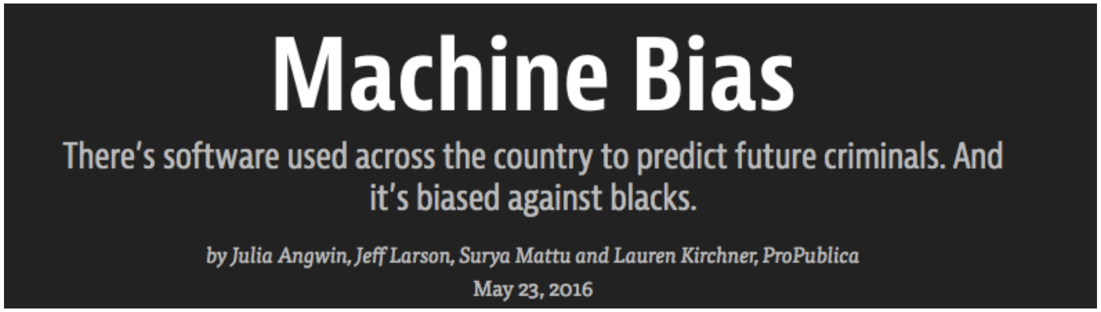
Kyle Behm accused Kroger and six other companies of discrimination against the mentally ill through their use of personality tests. *TROY STAINS FOR THE WALL STREET JOURNAL*

The Equal Employment Opportunity commission is **investigating whether personality tests discriminate against people with disabilities**.

As part of the investigation, officials are trying to determine if the tests **shut out people suffering from mental illnesses** such as depression or bipolar disorder, even if they have the right skills for the job.

http://www.wsj.com/articles/are-workplace-personality-tests-fair-1412044257

Racial bias in criminal sentencing



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

A commercial tool COMPAS automatically predicts some categories of future crime to assist in bail and sentencing decisions. It is used in courts in the US.

Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

The tool correctly predicts recidivism **61% of the time.**

Blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend.

The tool makes **the opposite mistake among whites**: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes.

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Transparency

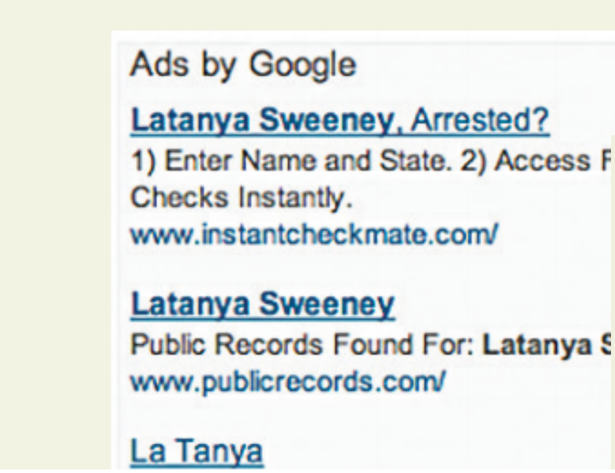
Transparency themes

- Online ad targeting: identifying the problem
  - Instant Checkmate / racially identifiable names
  - Ad Fisher
- Explaining black-box models (classifiers)
  - LIME: local interpretable explanations
  - QII: causal influence of features on outcomes
- Software design and testing for fairness (won't cover today)
- From auditing to interpretability: nutritional labels

Racially identifying names

[Latanya Sweeney; *CACM 2013*]

Google AdSense

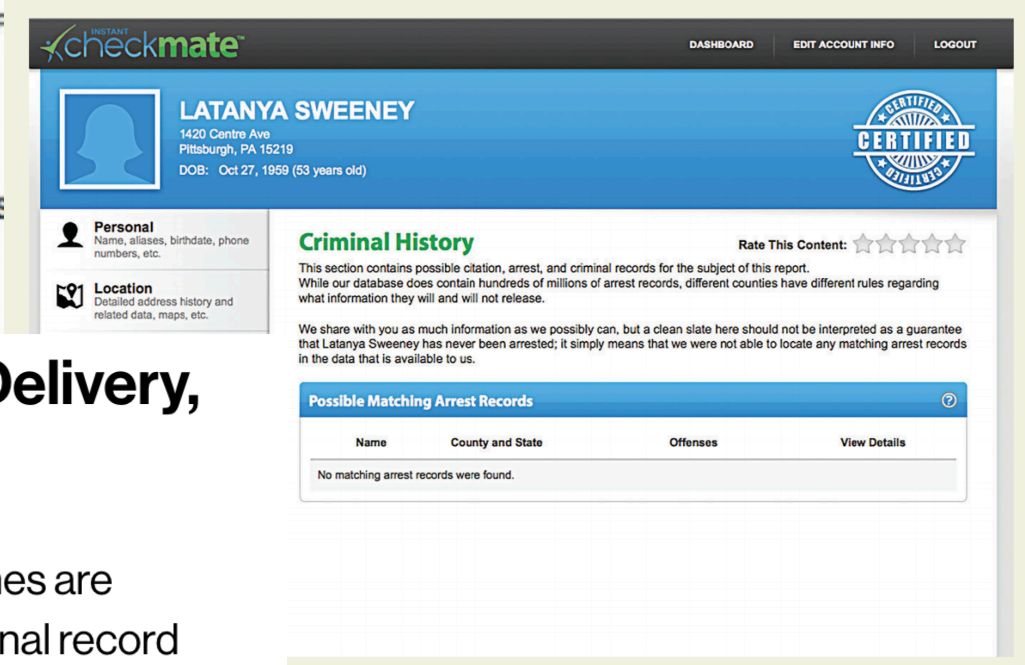


Ads by Google

Latanya Sweeney, Arrested?
1) Enter Name and State. 2) Access F
Checks Instantly.
www.instantcheckmate.com/

Latanya Sweeney
Public Records Found For: Latanya S
www.publicrecords.com/

La Tanya



LATANYA SWEENEY
1420 Centre Ave
Pittsburgh, PA 15219
DOB: Oct 27, 1959 (53 years old)

Personal Name, aliases, birthdate, phone numbers, etc.

Location Detailed address history and related data, maps, etc.

Criminal History

Rate This Content: ☆☆☆☆☆

This section contains possible citation, arrest, and criminal records for the subject of this report.
While our database does contain hundreds of millions of arrest records, different counties have different rules regarding what information they will and will not release.

We share with you as much information as we possibly can, but a clean slate here should not be interpreted as a guarantee that Latanya Sweeney has never been arrested; it simply means that we were not able to locate any matching arrest records in the data that is available to us.

Possible Matching Arrest Records

Name	County and State	Offenses	View Details
No matching arrest records were found.			

Racism is Poisoning Online Ad Delivery, Says Harvard Professor

Google searches involving black-sounding names are more likely to serve up ads suggestive of a criminal record than white-sounding names, says computer scientist

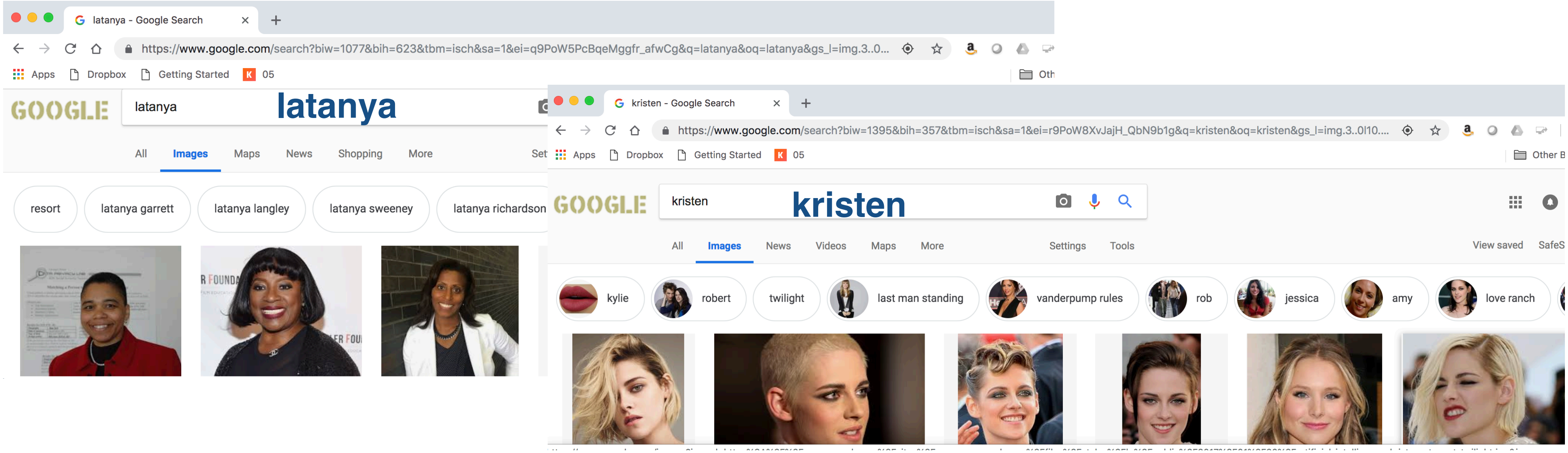
racially identifying names trigger ads suggestive of a criminal record

https://www.technologyreview.com/s/510646/racism-is-poisoning-online-ad-delivery-says-harvard-professor/

Racially identifying names: observations

[Latanya Sweeney; *CACM 2013*]

- Ads suggestive of a criminal record, linking to Instant Checkmate, appear on google.com and reuters.com in response to searchers for “Latanya Sweeney”, “Latanya Farrell”and “Latanya Lockett”*
- No Instant Checkmate ads when searching for “Kristen Haring”, “Kristen Sparrow”* and “Kristen Lindquist”*
- * next to a name that is associated with an arrest record



latanya - Google Search

https://www.google.com/search?biw=1077&bih=623&tbm=isch&sa=1&ei=q9PoW5PcBqeMggfr_afwCg&q=latanya&oq=latanya&gs_l=img.3..0...

kristen - Google Search

https://www.google.com/search?biw=1395&bih=357&tbm=isch&sa=1&ei=r9PoW8XvJajH_QbN9b1g&q=kristen&oq=kristen&gs_l=img.3..0l10....

Racially identifying names: details

[Latanya Sweeney; *CACM 2013*]

”A greater percentage of Instant Checkmate ads having the word arrest in ad text appeared for black-identifying first names than for white-identifying first names within professional and netizen subsets, too. On Reuters.com, which hosts Google AdSense ads, **a black-identifying name was 25% more likely to generate an ad suggestive of an arrest record**.”

More than 1,100 Instant Checkmate ads appeared on Reuters.com, with 488 having black-identifying first names; of these, 60% used arrest in the ad text. Of the 638 ads displayed with white-identifying names, 48% used arrest. This difference is statistically significant, with less than a 0.1% probability that the data can be explained by chance (chi-square test: X^2 (1)=14.32, p < 0.001).

The EEOC’s and U.S. Department of Labor’s adverse impact test for measuring discrimination is 77 in this case, so if this were an employment situation, a charge of discrimination might result. (The adverse impact test uses the ratio of neutral ads, or 100 minus the percentages given, to compute disparity: 100-60=40 and 100- 48=52; dividing 40 by 52 equals 77.)

Racially identifying names: Why?

[Latanya Sweeney; *CACM 2013*]

Possible explanations (from Latanya Sweeney):

- Does Instant Checkmate serve ads specifically for black-identifying names?
- Is Google's Adsense explicitly biased in this way?
- Does Google's Adsense learn racial bias based on from click-through rates?

How do we know which explanation is right?

We need transparency!

Response

https://www.technologyreview.com/s/510646/racism-is-poisoning-online-ad-delivery-says-harvard-professor/

In response to this blog post, a **Google** spokesperson sends the following comment:

“AdWords does not conduct any racial profiling. We also have an “anti” and violence policy which states that we will not allow ads that advocate against an organisation, person or group of people. It is up to individual advertisers to decide which keywords they want to choose to trigger their ads.”

Instantcheckmate.com sends the following statement:

“As a point of fact, Instant Checkmate would like to state unequivocally that it has never engaged in racial profiling in Google AdWords. We have absolutely no technology in place to even connect a name with a race and have never made any attempt to do so. The very idea is contrary to our company’s most deeply held principles and values.”

Who is responsible?

- Who benefits?
- Who is harmed?
- What does the law say?
- Who is in a position to mitigate?

transparency …. responsibility …. TRUST

Detour: Barrow, Alaska, 1979

Native leaders and city officials, worried about drinking and associated violence in their community **invited a group of sociology researchers** to assess the problem and work with them to devise solutions.

Methodology:

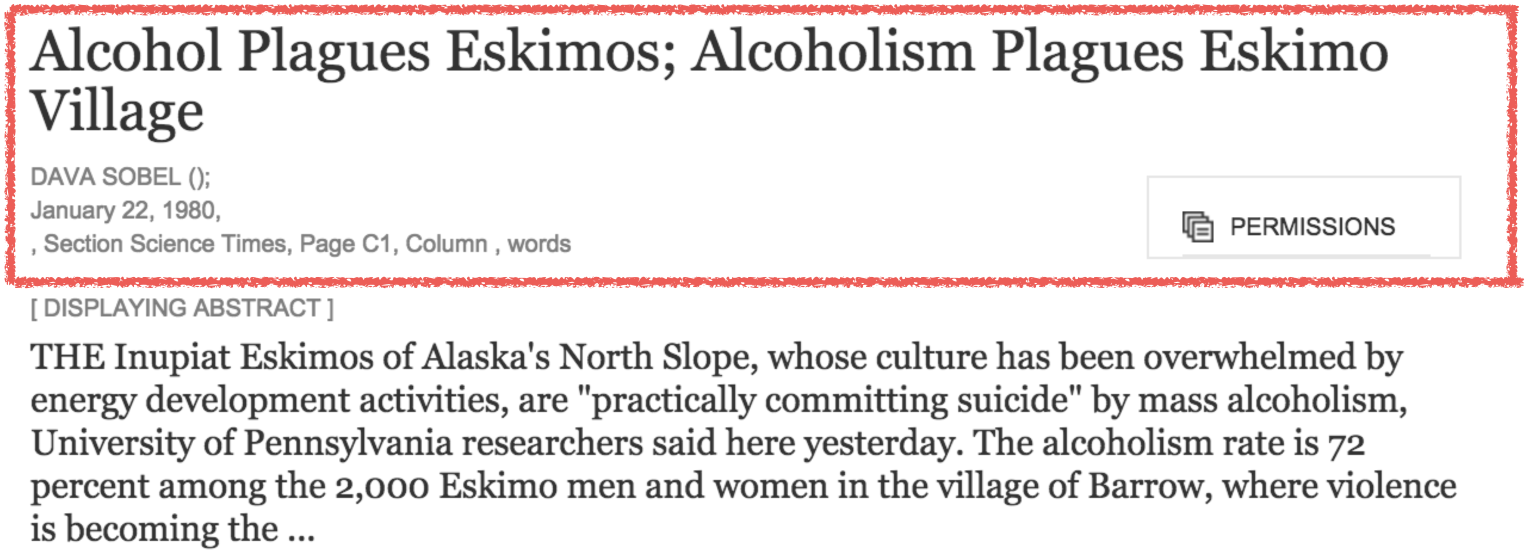
- 10% representative sample (N=88) of everyone over the age of 15 using a 1972 demographic survey
- Interviewed on attitudes and values about use of alcohol
- Obtained psychological histories & drinking behavior
- Given the Michigan Alcoholism Screening Test
- Asked to draw a picture of a person (used to determine cultural identity)

Barrow
Nome
ALASKA
Anchorage
Gulf of Alaska
Google

Study “results”

At the conclusion of the study researchers formulated a report entitled **“The Inupiat, Economics and Alcohol on the Alaskan North Slope”**, released **simultaneously** at a press release and to the Barrow community.

The press release was picked up by the New York Times, who ran a front page story entitled **“Alcohol Plagues Eskimos”**



Alcohol Plagues Eskimos; Alcoholism Plagues Eskimo Village

DAVA SOBEL ();
January 22, 1980,
, Section Science Times, Page C1, Column , words

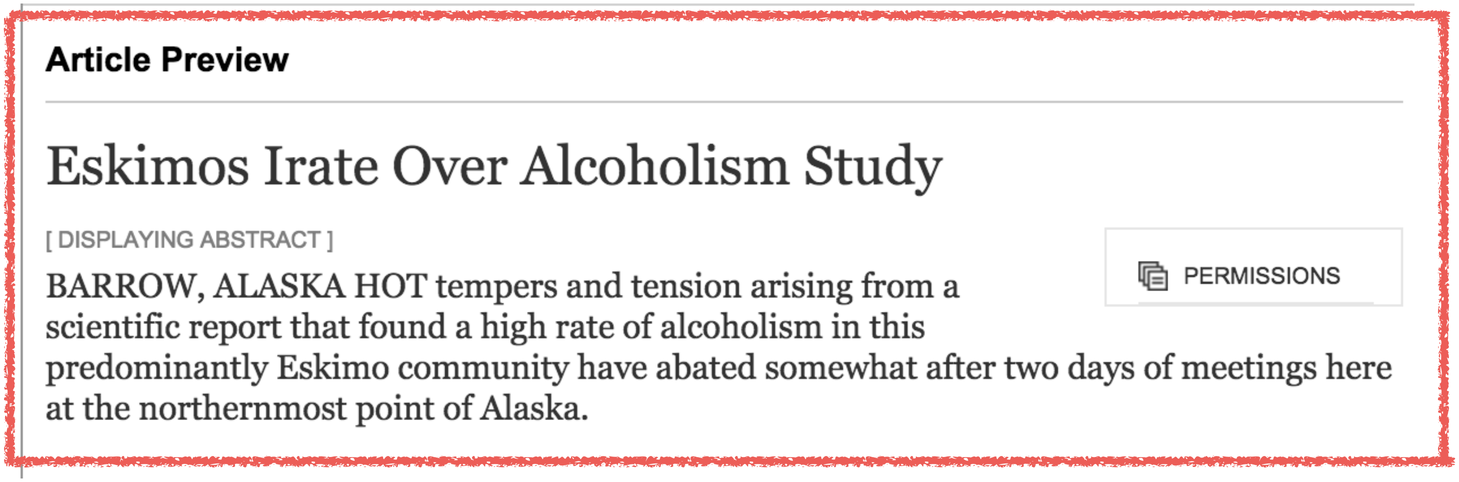
[DISPLAYING ABSTRACT]

THE Inupiat Eskimos of Alaska's North Slope, whose culture has been overwhelmed by energy development activities, are "practically committing suicide" by mass alcoholism, University of Pennsylvania researchers said here yesterday. The alcoholism rate is 72 percent among the 2,000 Eskimo men and women in the village of Barrow, where violence is becoming the ...

Harms and backlash

Study **results were revealed** in the context of a press conference that was held far from the Native village, and **without the presence, much less the knowledge or consent**, of any community member who might have been able to present any context concerning the socioeconomic conditions of the village.

Study results suggested that nearly all adults in the community were alcoholics. In addition to the shame felt by community members, the town's Standard and Poor bond rating suffered as a result, which in turn decreased the tribe's ability to secure funding for much needed projects.



Article Preview

Eskimos Irate Over Alcoholism Study

[DISPLAYING ABSTRACT]

BARROW, ALASKA HOT tempers and tension arising from a scientific report that found a high rate of alcoholism in this predominantly Eskimo community have abated somewhat after two days of meetings here at the northernmost point of Alaska.

Problems

Methodological

Edward F. Foulks, M.D., "Misalliances In The Barrow Alcohol Study"

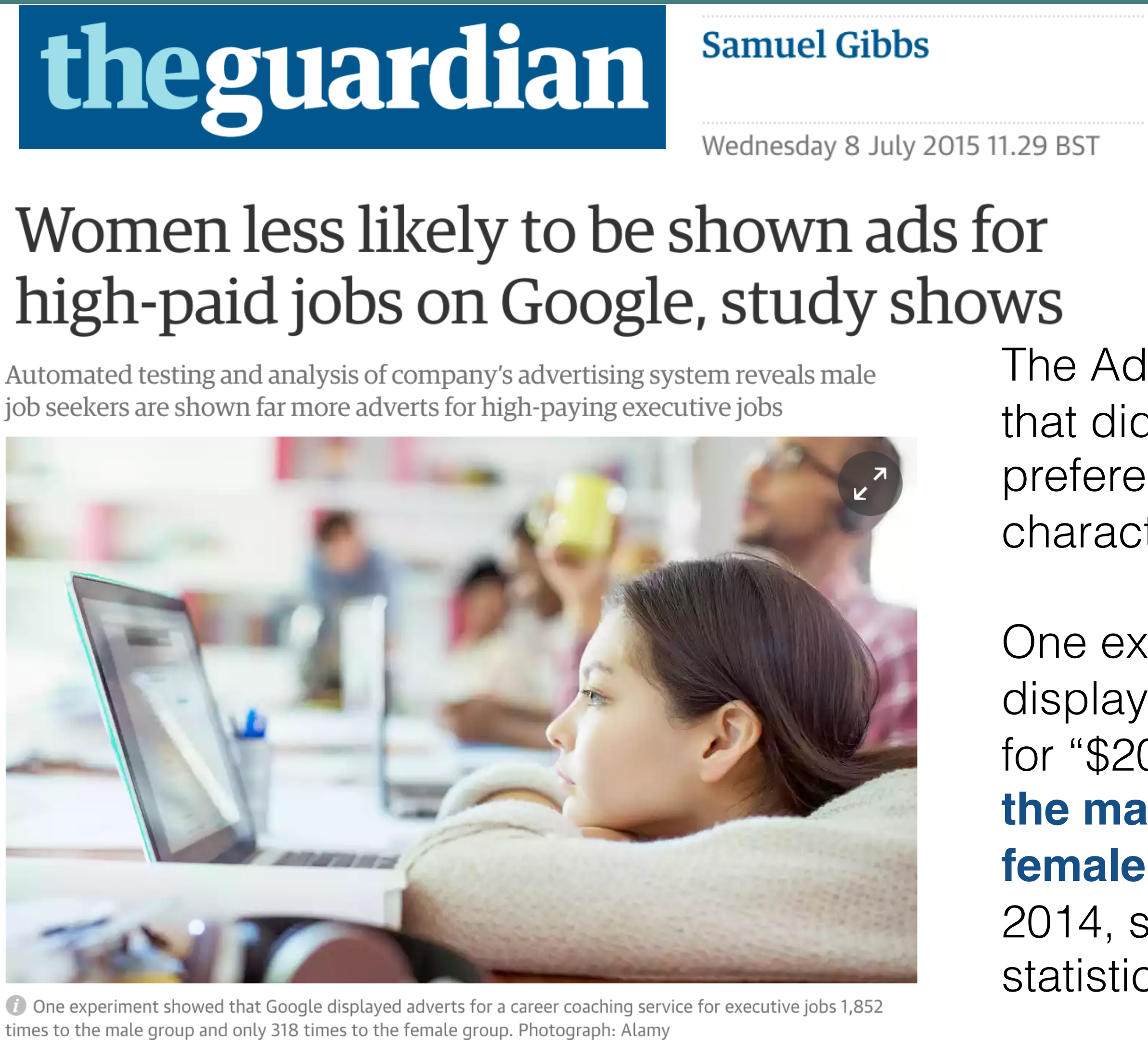
- "The authors once again met with the Barrow Technical Advisory Group, who stated their concern that only Natives were studied, and that outsiders in town had not been included."
- "The estimates of the frequency of intoxication based on association with the probability of being detained were termed "ludicrous, both logically and statistically."

Ethical

- Participants not in control of their data
- Significant harm: social (stigmatization) and financial (bond rating)
- No laws were broken, and harms are not about individual privacy!
- **Who benefits? Who is harmed?**

data protection …. responsibility …. TRUST

Back to the future: Online job ads



theguardian

Samuel Gibbs

Wednesday 8 July 2015 11.29 BST

Women less likely to be shown ads for high-paid jobs on Google, study shows

Automated testing and analysis of company's advertising system reveals male job seekers are shown far more adverts for high-paying executive jobs

One experiment showed that Google displayed adverts for a career coaching service for executive jobs 1,852 times to the male group and only 318 times to the female group. Photograph: Alamy

The AdFisher tool simulated job seekers that did not differ in browsing behavior, preferences or demographic characteristics, except in gender.

One experiment showed that Google displayed ads for a career coaching service for "$200k+" executive jobs **1,852 times to the male group and only 318 times to the female group**. Another experiment, in July 2014, showed a similar trend but was not statistically significant.

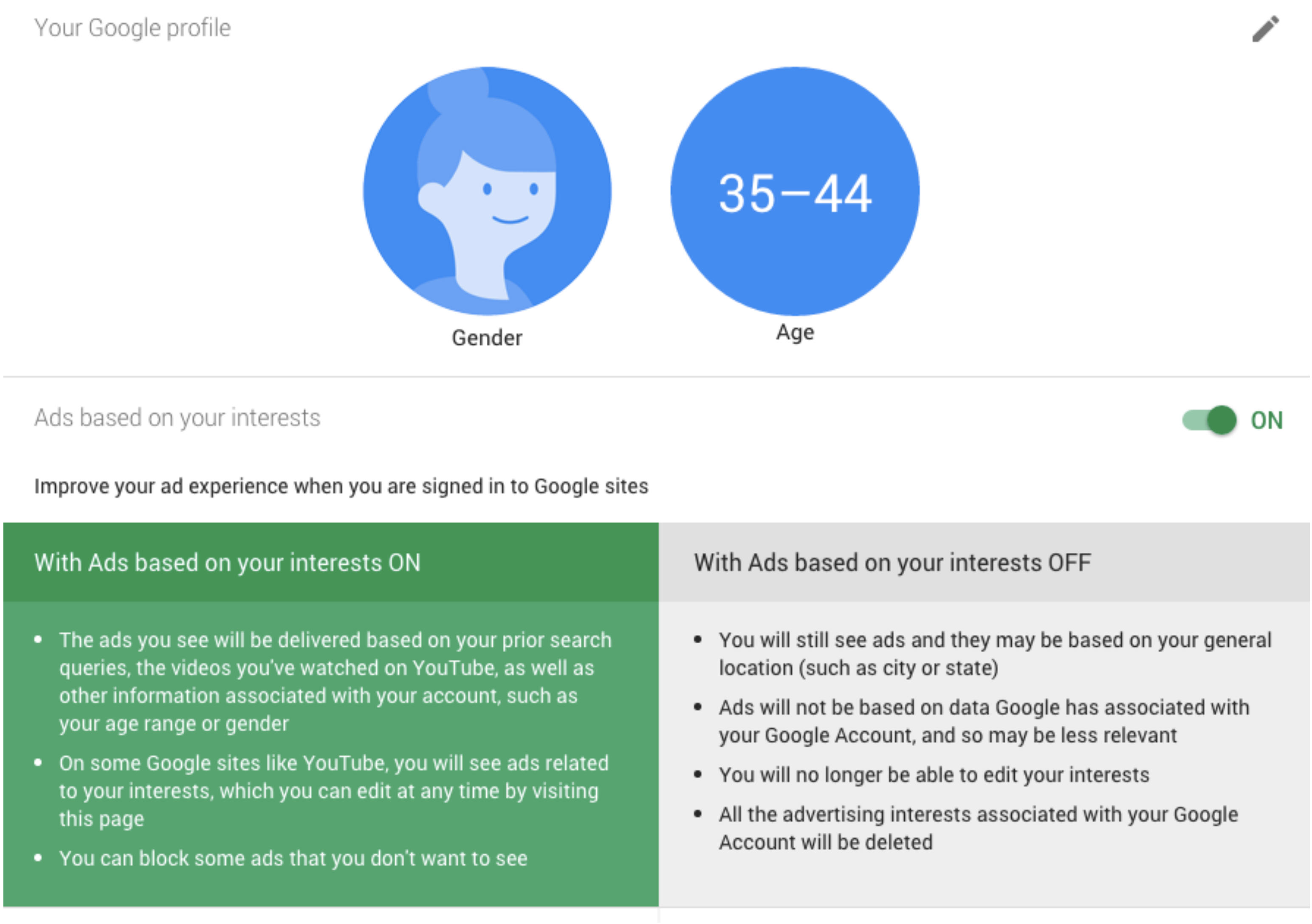
https://www.theguardian.com/technology/2015/jul/08/women-less-likely-ads-high-paid-jobs-google-study

Ad targeting online

- **Users** browse the Web, consume content, consume ads (see / click / purchase)
- **Content providers** outsource advertising to third-party ad networks, e.g., Google's DoubleClick
- **Ad networks** track users across sites, to get a global view of users' behaviors
- **Google Ad Settings** aims to provide **transparency** / give **control to users** over the ads that they see

do users truly have transparency / choice or is this a placebo button?

Google Ads Settings



Your Google profile

Gender

35–44

Age

Ads based on your interests ON

Improve your ad experience when you are signed in to Google sites

With Ads based on your interests ON

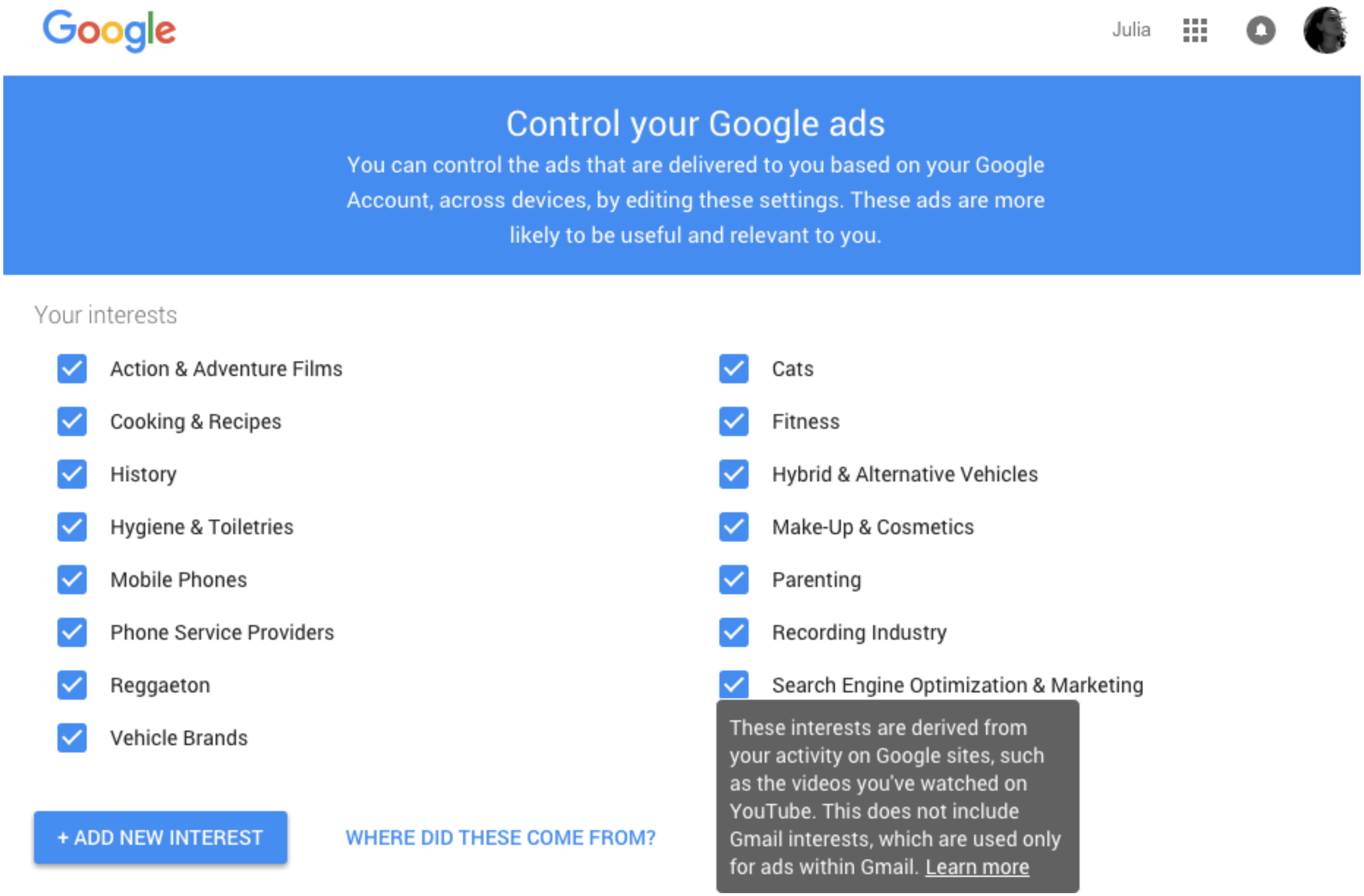
- The ads you see will be delivered based on your prior search queries, the videos you've watched on YouTube, as well as other information associated with your account, such as your age range or gender
- On some Google sites like YouTube, you will see ads related to your interests, which you can edit at any time by visiting this page
- You can block some ads that you don't want to see

With Ads based on your interests OFF

- You will still see ads and they may be based on your general location (such as city or state)
- Ads will not be based on data Google has associated with your Google Account, and so may be less relevant
- You will no longer be able to edit your interests
- All the advertising interests associated with your Google Account will be deleted

http://www.google.com/settings/ads

Google Ads Settings



Control your Google ads

You can control the ads that are delivered to you based on your Google Account, across devices, by editing these settings. These ads are more likely to be useful and relevant to you.

Your interests

- ☑ Action & Adventure Films
- ☑ Cooking & Recipes
- ☑ History
- ☑ Hygiene & Toiletries
- ☑ Mobile Phones
- ☑ Phone Service Providers
- ☑ Reggaeton
- ☑ Vehicle Brands
- ☑ Cats
- ☑ Fitness
- ☑ Hybrid & Alternative Vehicles
- ☑ Make-Up & Cosmetics
- ☑ Parenting
- ☑ Recording Industry
- ☑ Search Engine Optimization & Marketing

These interests are derived from your activity on Google sites, such as the videos you've watched on YouTube. This does not include Gmail interests, which are used only for ads within Gmail. Learn more

http://www.google.com/settings/ads

AdFisher

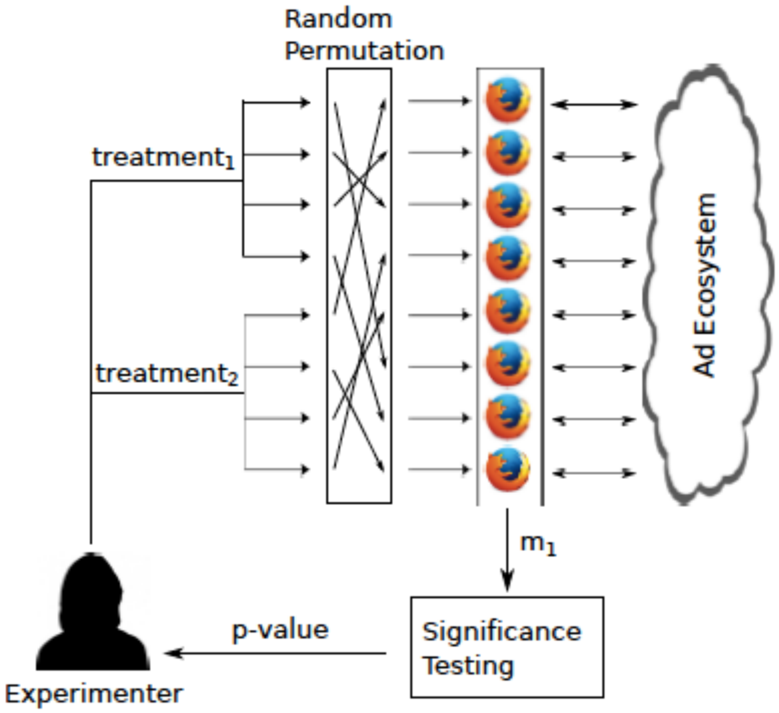
[A. Datta, M. Tschantz, A. Datta; *PETS 2015*]

From anecdotal evidence to statistical insight:

How do user behaviors, ads and ad settings interact?

Automated randomized controlled experiments for studying online tracking

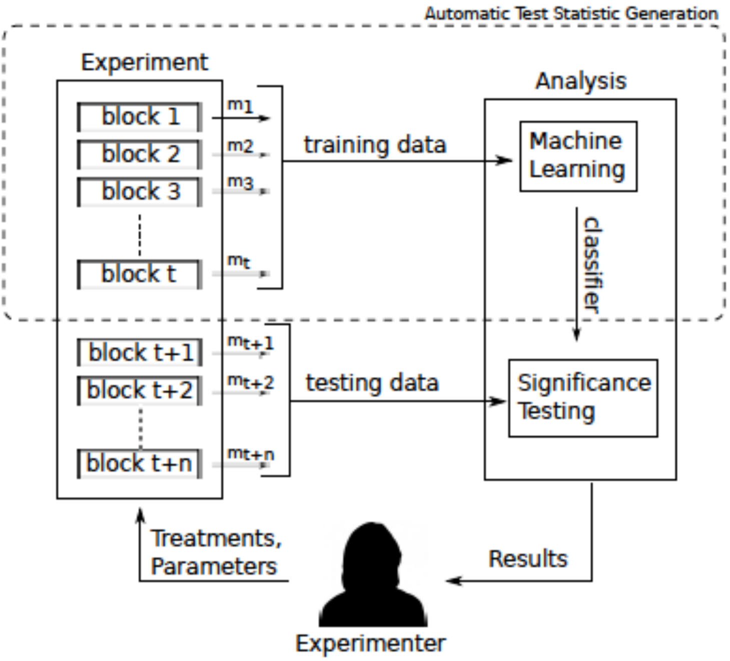
Individual data use transparency: ad network must share the information it uses about the user to select which ads to serve to him



AdFisher: methodology

[A. Datta, M. Tschantz, A. Datta; *PETS 2015*]

- Browser-based experiments, simulated users
  - **input**: (1) visits to content providing websites; (2) interactions with Google Ad Settings
  - **output**: (1) ads shown to users by Google; (2) change in Google Ad Settings
- Fisher randomized hypothesis testing
  - **null hypothesis** inputs do not affect outputs
  - control and experimental treatments
  - AdFisher can help select a test statistic



AdFisher: gender and jobs

[A. Datta, M. Tschantz, A. Datta; *PETS 2015*]

Non-discrimination: Users differing only in protected attributes are treated similarly

Causal test: Find that a protected attribute changes ads

Experiment 1: **gender and jobs**

Specify gender (male/female) in Ad Settings, simulate interest in jobs by visiting employment sites, collect ads from Times of India or the Guardian

Result: males were shown ads for higher-paying jobs significantly more often than females (1852 vs. 318)

violation

AdFisher: substance abuse

[A. Datta, M. Tschantz, A. Datta; *PETS 2015*]

Transparency: User can view data about him used for ad selection

Causal test: Find attribute that changes ads but not settings

Experiment 2: **substance abuse**

Simulate interest in substance abuse in the experimental group but not in the control group, check for differences in Ad Settings, collect ads from Times of India

Result: no difference in Ad Settings between the groups, yet significant differences in what ads are served: rehab vs. stocks + driving jobs

violation

AdFisher: online dating

[A. Datta, M. Tschantz, A. Datta; *PETS 2015*]

Ad choice: Removing an interest decreases the number of ads related to that interest.

Causal test: Find that removing an interest causes a decrease in related ads

Experiment 3: **online dating**

Simulate interest in online dating in both groups, remove “Dating & Personals” from the interests on Ad Settings for experimental group, collect ads

Result: members of experimental group do not get ads related to dating, while members of the control group do

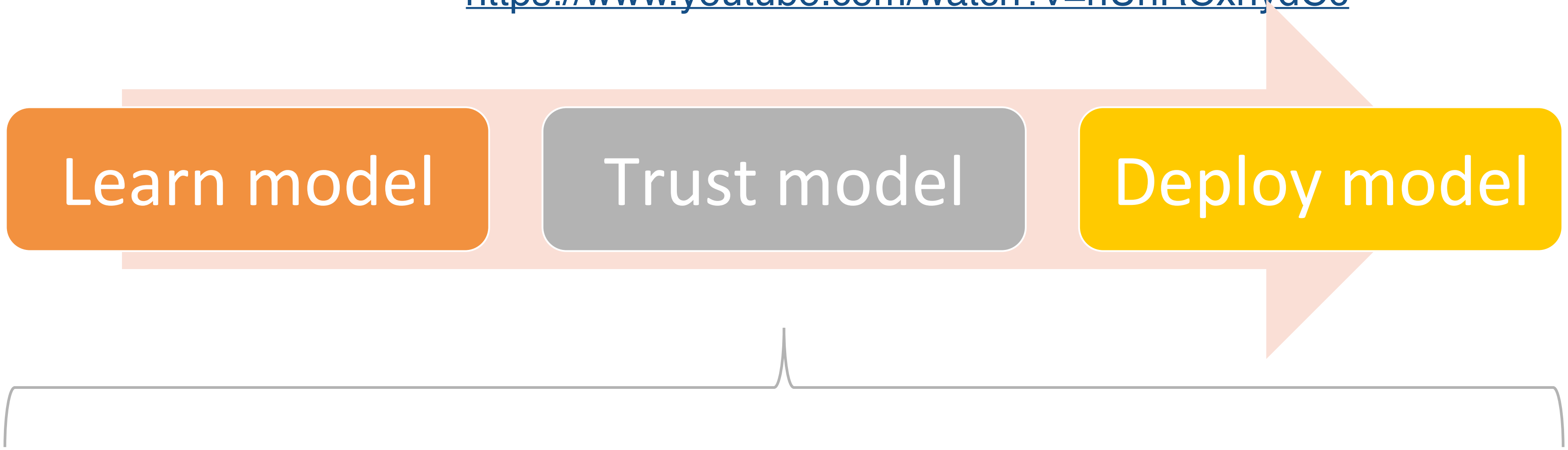
compliance

Explaining black-box classifiers

LIME: Local explanations of classifiers

[M. T. Ribeiro, S. Singh, C. Guestrin; *KDD 2016*]

https://www.youtube.com/watch?v=hUnRCxnydCc



Learn model

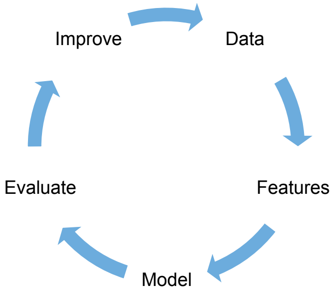
Trust model

Deploy model

Trust AI system

Make better decisions

Improve model



slide by Marco Tulio Ribeiro, KDD 2016

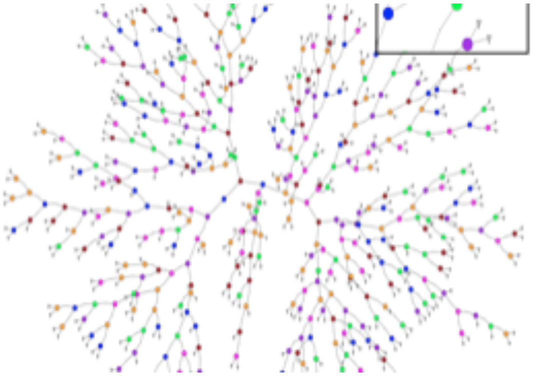
LIME: Local explanations of classifiers

[M. T. Ribeiro, S. Singh, C. Guestrin; *KDD 2016*]

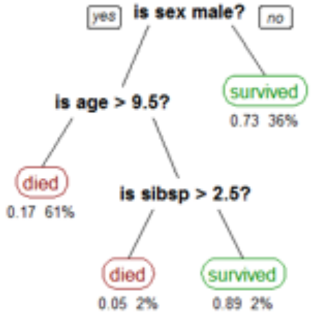
Three must-haves for a good explanation

Interpretable

- Humans can easily interpret reasoning



Definitely
not interpretable



Potentially
interpretable

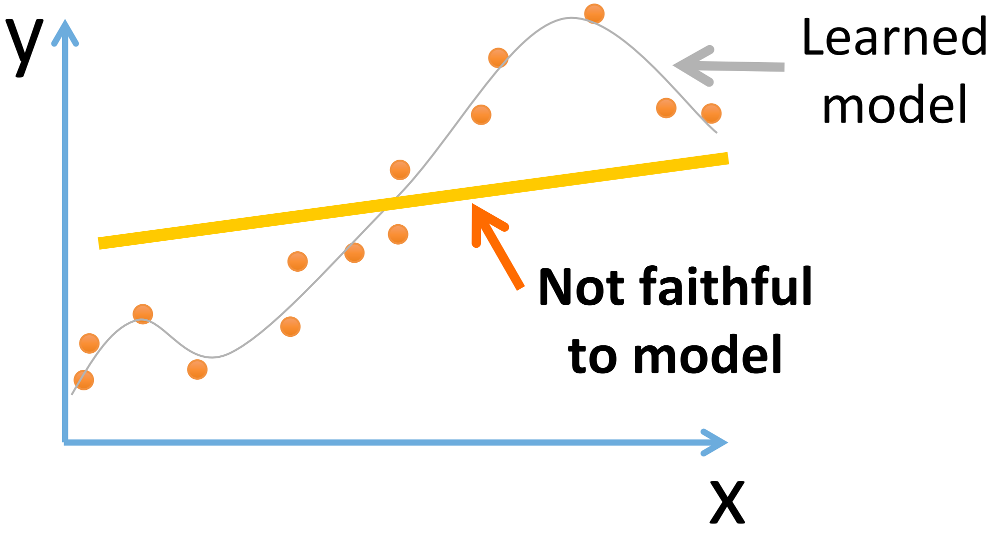
slide by Marco Tulio Ribeiro, KDD 2016

LIME: Local explanations of classifiers

[M. T. Ribeiro, S. Singh, C. Guestrin; *KDD 2016*]

Three must-haves for a good explanation

Interpretable	• Humans can easily interpret reasoning
Faithful	• Describes how this model actually behaves



slide by Marco Tulio Ribeiro, KDD 2016

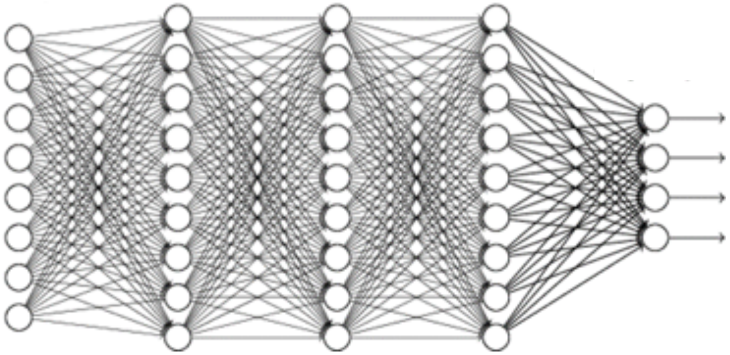
LIME: Local explanations of classifiers

[M. T. Ribeiro, S. Singh, C. Guestrin; *KDD 2016*]

Three must-haves for a good explanation

Interpretable	• Humans can easily interpret reasoning
Faithful	• Describes how this model actually behaves
Model agnostic	• Can be used for *any* ML model

Can explain this mess ☺

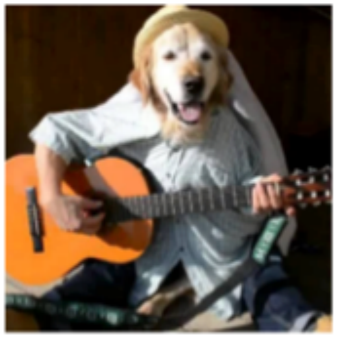


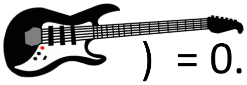
slide by Marco Tulio Ribeiro, KDD 2016

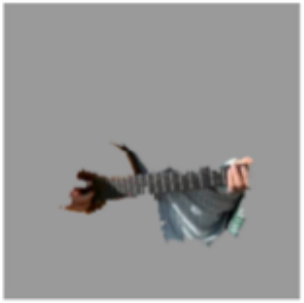
Local explanations of classifiers

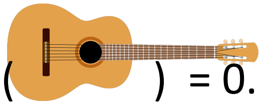
[M. T. Ribeiro, S. Singh, C. Guestrin; *KDD 2016*]

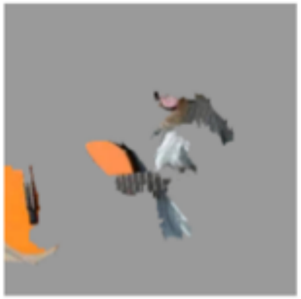
Explaining Google's Inception NN



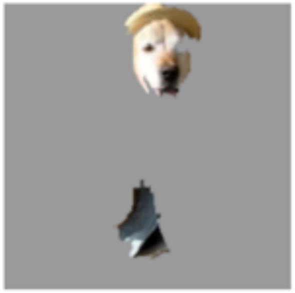
P() = 0.32



P() = 0.24



P( ) = 0.21

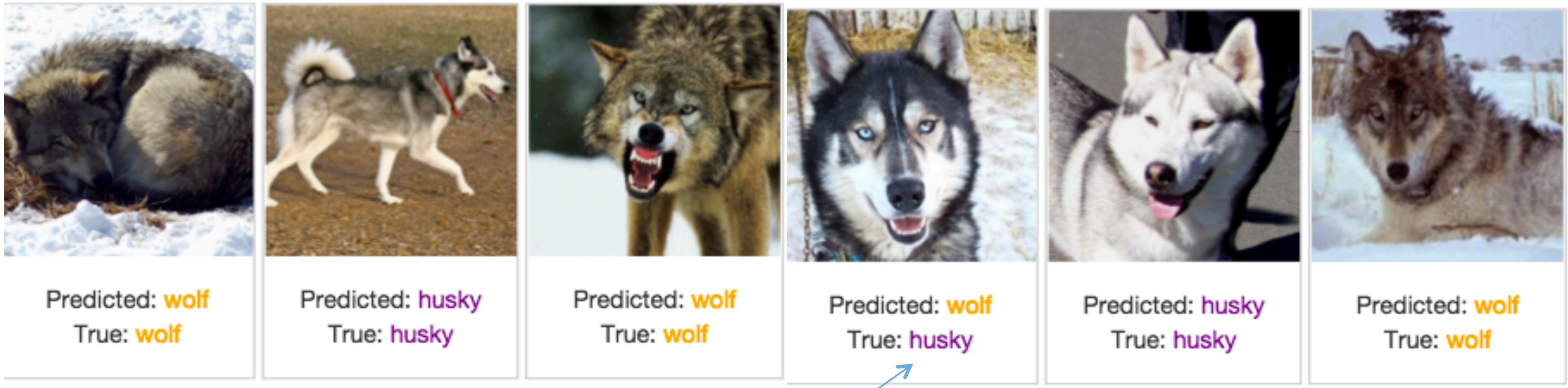


slide by Marco Tulio Ribeiro, KDD 2016

Local explanations of classifiers

[M. T. Ribeiro, S. Singh, C. Guestrin; *KDD 2016*]

Train a neural network to predict wolf v. husky



Only 1 mistake!!!

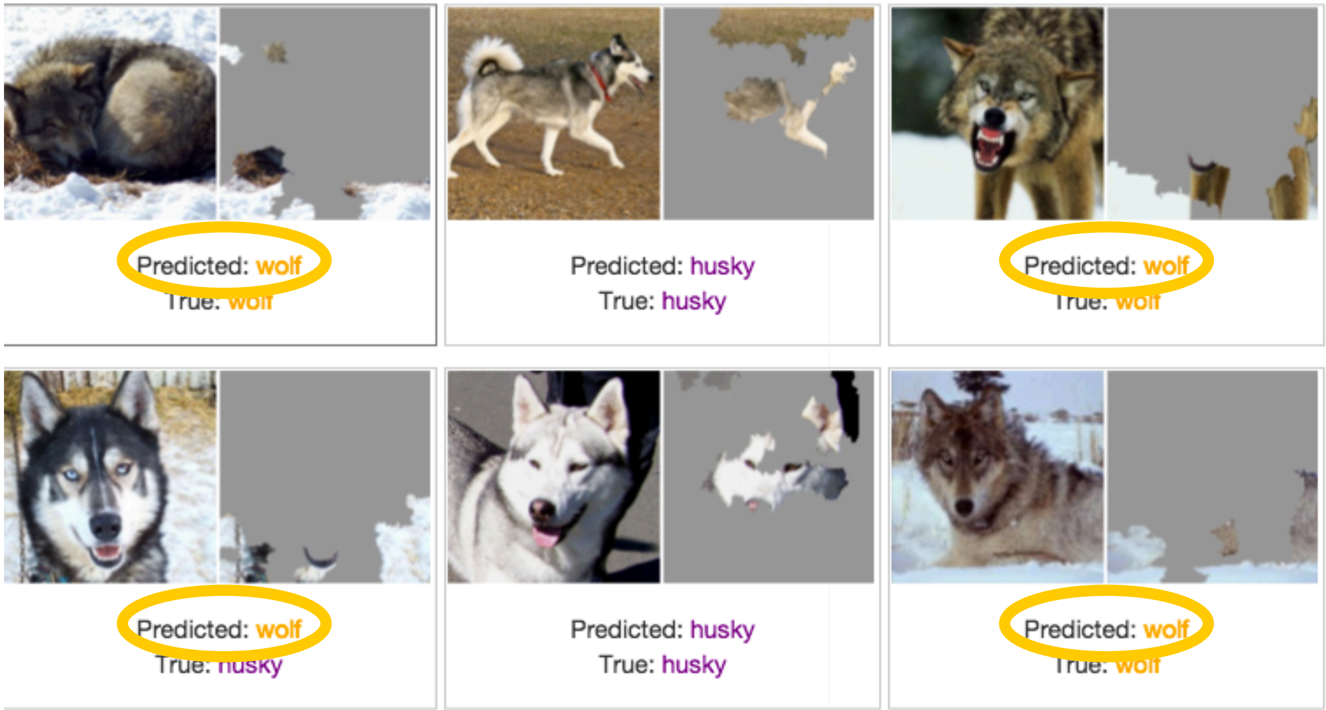
Do you trust this model?
How does it distinguish between huskies and wolves?

slide by Marco Tulio Ribeiro, KDD 2016

Local explanations of classifiers

[M. T. Ribeiro, S. Singh, C. Guestrin; *KDD 2016*]

Explanations for neural network prediction



We've built a great snow detector… ☹

slide by Marco Tulio Ribeiro, KDD 2016

Nutritional labels for data and models

Transparency in ranking

Input: database of items (individuals, colleges, cars, …)

Score-based ranker: computes the score of each item using a **known formula**, e.g., monotone aggregation, then sorts items on score

Output: permutation of the items (complete or top-k)

Do we have transparency?

We have syntactic transparency, but lack interpretability!

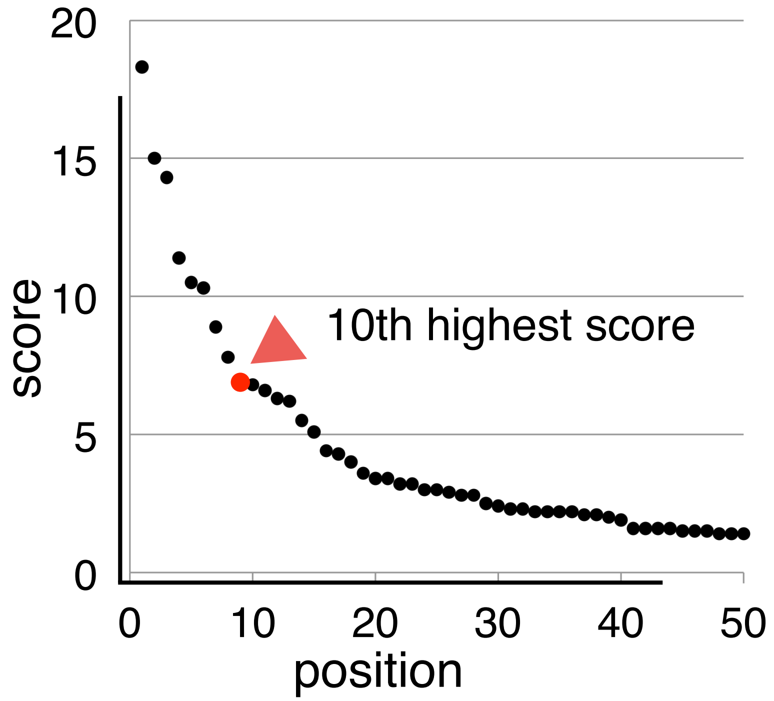
https://freedom-to-tinker.com/2018/05/03/refining-the-concept-of-a-nutritional-label-for-data-and-models/

https://freedom-to-tinker.com/2016/08/05/revealing-algorithmic-rankers/

Opacity in algorithmic rankers

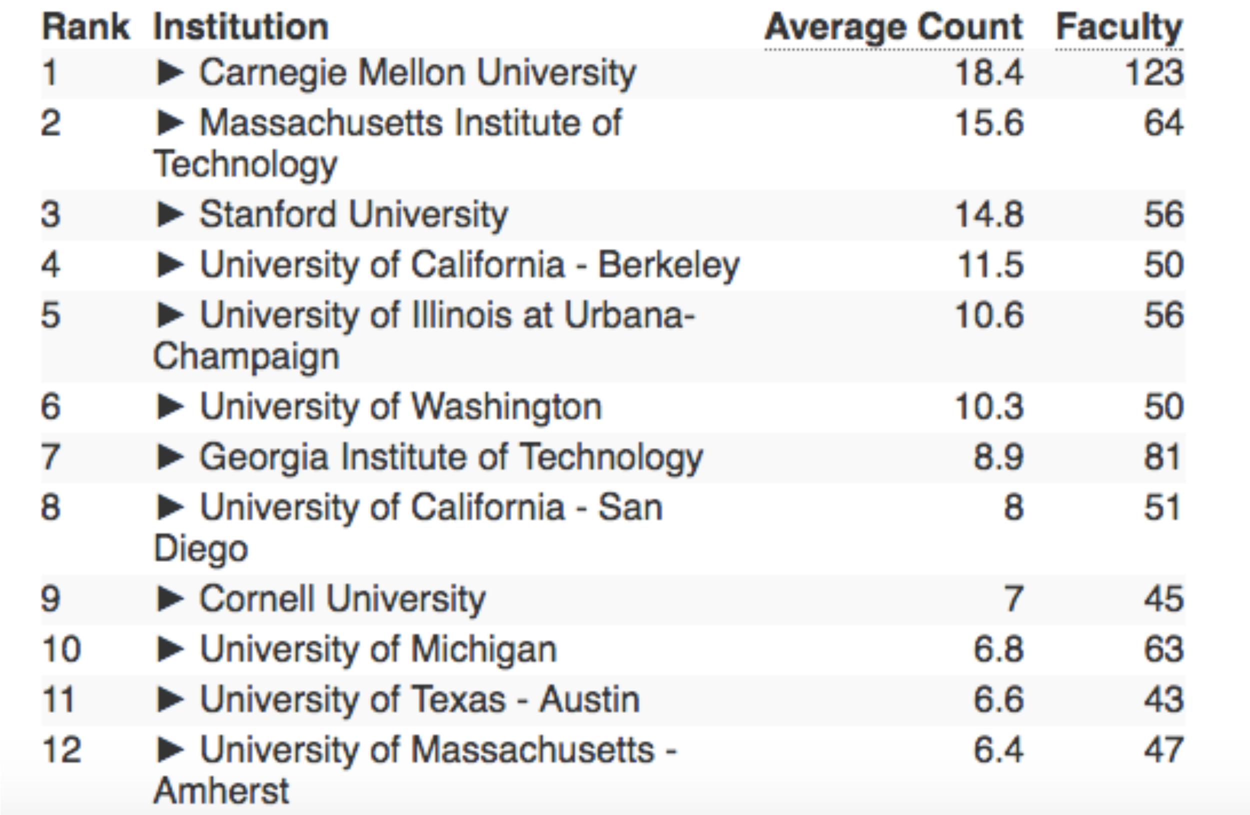
Reason 1: The scoring formula alone does not indicate the relative rank of an item.

Scores are absolute, rankings are relative. Is 5 a good score? What about 10? 15?



Opacity in algorithmic rankers

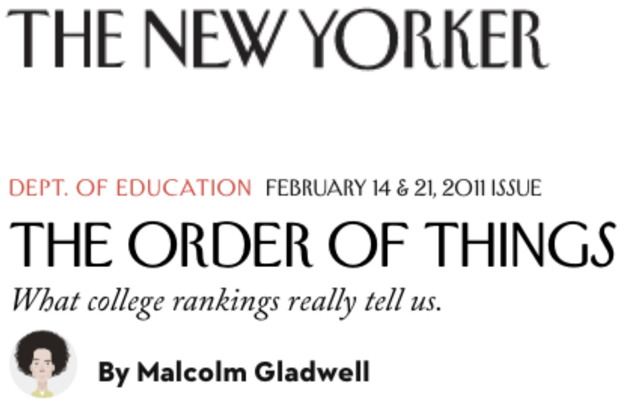
Reason 2: A ranking may be unstable if there are tied or nearly-tied items.



Rank	Institution	Average Count	Faculty
1	▶ Carnegie Mellon University	18.4	123
2	▶ Massachusetts Institute of Technology	15.6	64
3	▶ Stanford University	14.8	56
4	▶ University of California - Berkeley	11.5	50
5	▶ University of Illinois at Urbana-Champaign	10.6	56
6	▶ University of Washington	10.3	50
7	▶ Georgia Institute of Technology	8.9	81
8	▶ University of California - San Diego	8	51
9	▶ Cornell University	7	45
10	▶ University of Michigan	6.8	63
11	▶ University of Texas - Austin	6.6	43
12	▶ University of Massachusetts - Amherst	6.4	47

Opacity in algorithmic rankers

Reason 3: A ranking methodology may be unstable: small changes in weights can trigger significant re-shuffling.



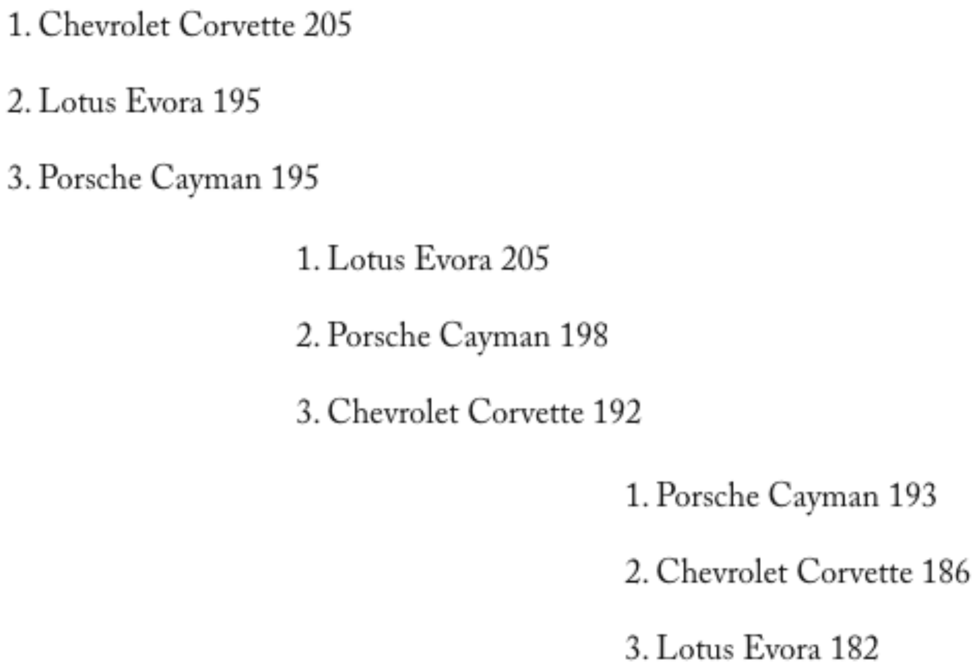
THE NEW YORKER

DEPT. OF EDUCATION FEBRUARY 14 & 21, 2011 ISSUE

THE ORDER OF THINGS

What college rankings really tell us.

By Malcolm Gladwell



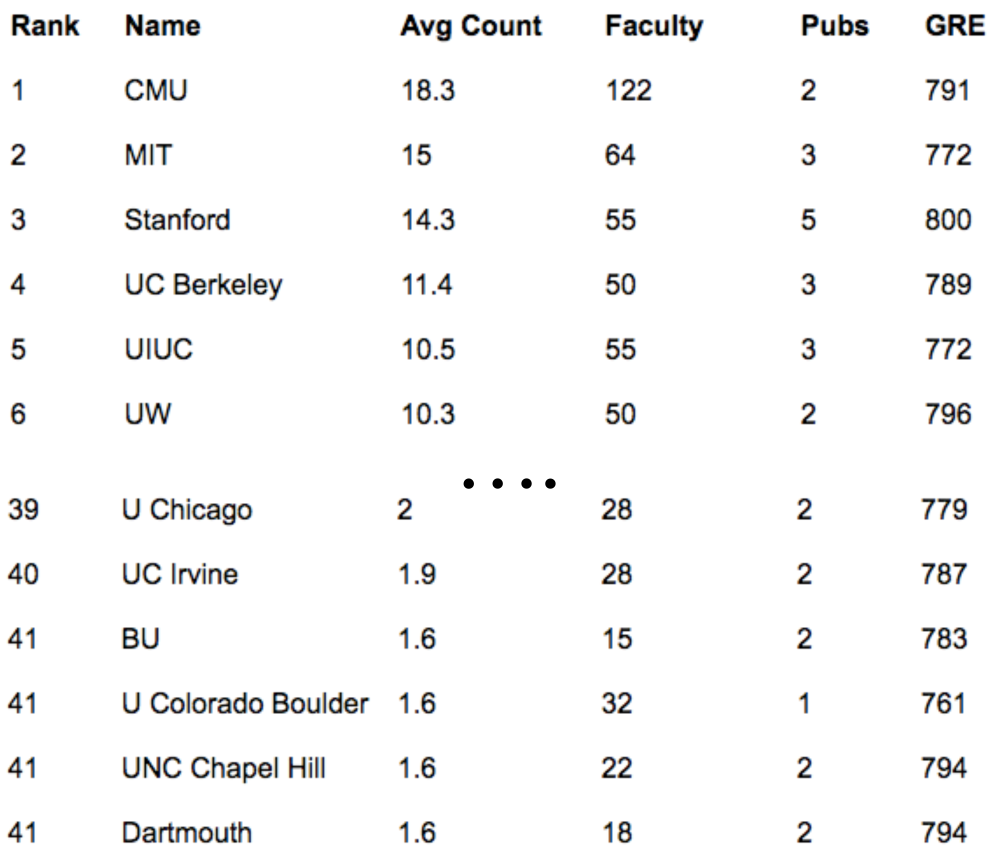
1. Chevrolet Corvette 205
2. Lotus Evora 195
3. Porsche Cayman 195

1. Lotus Evora 205
2. Porsche Cayman 198
3. Chevrolet Corvette 192

1. Porsche Cayman 193
2. Chevrolet Corvette 186
3. Lotus Evora 182

Opacity in algorithmic rankers

Reason 4: The **weight of an attribute** in the scoring formula **does not determine its impact** on the outcome.



Rank	Name	Avg Count	Faculty	Pubs	GRE
1	CMU	18.3	122	2	791
2	MIT	15	64	3	772
3	Stanford	14.3	55	5	800
4	UC Berkeley	11.4	50	3	789
5	UIUC	10.5	55	3	772
6	UW	10.3	50	2	796
		….			
39	U Chicago	2	28	2	779
40	UC Irvine	1.9	28	2	787
41	BU	1.6	15	2	783
41	U Colorado Boulder	1.6	32	1	761
41	UNC Chapel Hill	1.6	22	2	794
41	Dartmouth	1.6	18	2	794

Given a score function:

$$0.2 * faculty +$$
$$0.3 * avg\ cnt +$$
$$0.5 * gre$$

Rankings are not benign!

THE NEW YORKER

DEPT. OF EDUCATION FEBRUARY 14 & 21, 2011 ISSUE

THE ORDER OF THINGS

What college rankings really tell us.

By Malcolm Gladwell

Rankings are not benign. They enshrine very particular ideologies, and, at a time when American higher education is facing a crisis of accessibility and affordability, we have adopted a de-facto standard of college quality that is uninterested in both of those factors. And why? Because a group of magazine analysts in an office building in Washington, D.C., decided twenty years ago to value selectivity over efficacy, to use proxies that scarcely relate to what they're meant to be proxies for, and to pretend that they can compare a large, diverse, low-cost land-grant university in rural Pennsylvania with a small, expensive, private Jewish university on two campuses in Manhattan.

Harms of opacity

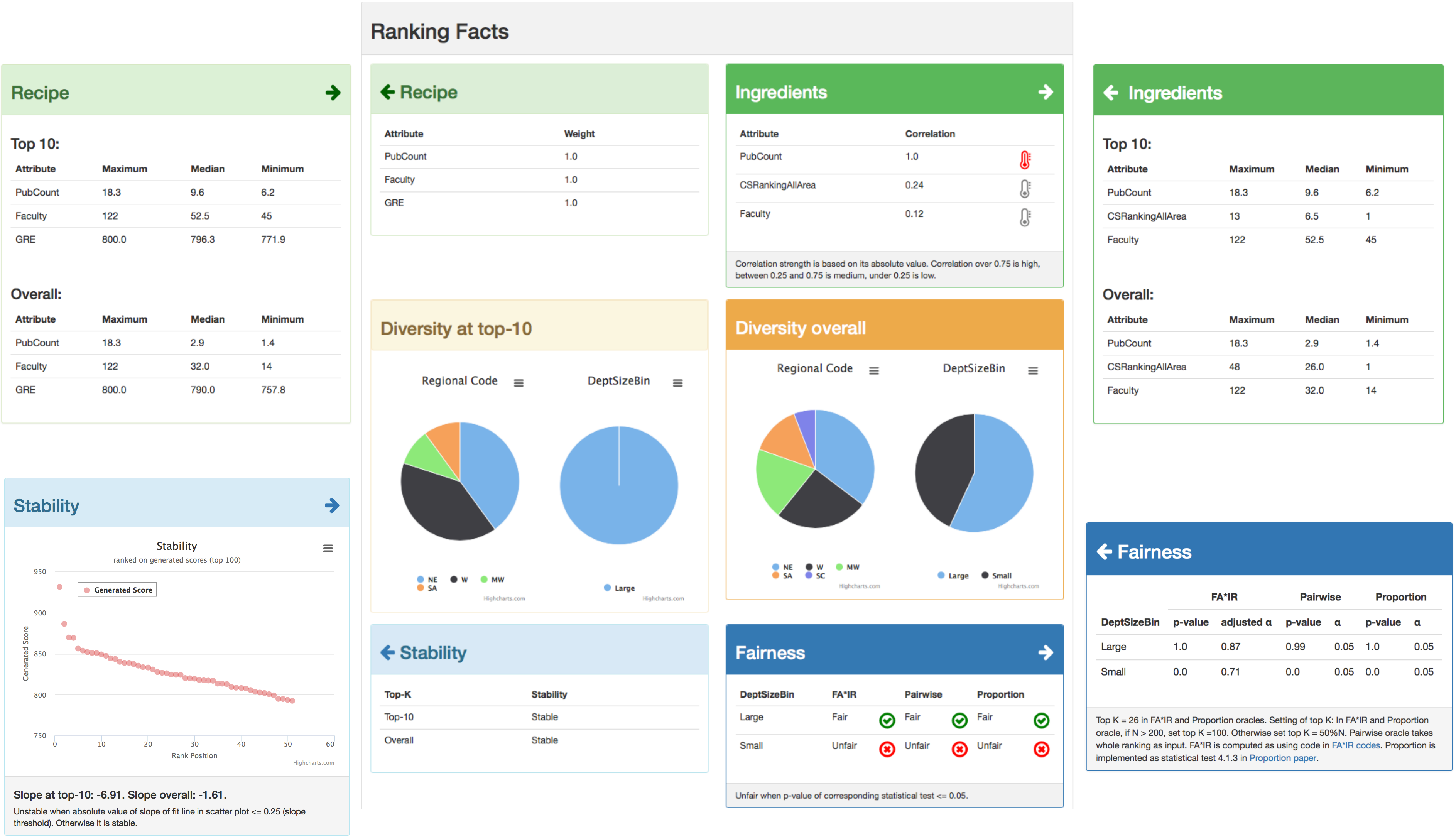
1. **Due process / fairness.** The subjects of the ranking cannot have confidence that their ranking is meaningful or correct, or that they have been treated like similarly situated subjects - *procedural regularity*

2. **Hidden normative commitments.** What factors does the vendor encode in the scoring ranking process (syntactically)? What are the *actual* effects of the scoring / ranking process? Is it stable? How was it validated?

Harms of opacity

3. **Interpretability.** Especially where ranking algorithms are performing a public function, **political legitimacy** requires that the public be able to interpret algorithmic outcomes in a meaningful way. Avoid *algocracy*: the rule by incontestable algorithms.

4. **Meta-methodological assessment.** Is *a* ranking / *this* ranking appropriate here? Can we use a process if it cannot be explained? Probably yes, for recommending movies; probably not for college admissions.



Ranking Facts

Recipe

Top 10:

Attribute	Maximum	Median	Minimum
PubCount	18.3	9.6	6.2
Faculty	122	52.5	45
GRE	800.0	796.3	771.9

Overall:

Attribute	Maximum	Median	Minimum
PubCount	18.3	2.9	1.4
Faculty	122	32.0	14
GRE	800.0	790.0	757.8

Stability

Slope at top-10: -6.91. Slope overall: -1.61.

Unstable when absolute value of slope of fit line in scatter plot <= 0.25 (slope threshold). Otherwise it is stable.

Recipe

Attribute	Weight
PubCount	1.0
Faculty	1.0
GRE	1.0

Ingredients

Attribute	Correlation
PubCount	1.0
CSRankingAllArea	0.24
Faculty	0.12

Correlation strength is based on its absolute value. Correlation over 0.75 is high, between 0.25 and 0.75 is medium, under 0.25 is low.

Diversity at top-10

Diversity overall

Stability

Top-K	Stability
Top-10	Stable
Overall	Stable

Fairness

DeptSizeBin	FA*IR		Pairwise		Proportion	
Large	Fair	✓	Fair	✓	Fair	✓
Small	Unfair	✗	Unfair	✗	Unfair	✗

Unfair when p-value of corresponding statistical test <= 0.05.

Ingredients

Top 10:

Attribute	Maximum	Median	Minimum
PubCount	18.3	9.6	6.2
CSRankingAllArea	13	6.5	1
Faculty	122	52.5	45

Overall:

Attribute	Maximum	Median	Minimum
PubCount	18.3	2.9	1.4
CSRankingAllArea	48	26.0	1
Faculty	122	32.0	14

Fairness

	FA*IR		Pairwise		Proportion	
DeptSizeBin	p-value	adjusted α	p-value	α	p-value	α
Large	1.0	0.87	0.99	0.05	1.0	0.05
Small	0.0	0.71	0.0	0.05	0.0	0.05

Top K = 26 in FA*IR and Proportion oracles. Setting of top K: In FA*IR and Proportion oracle, if N > 200, set top K =100. Otherwise set top K = 50%N. Pairwise oracle takes whole ranking as input. FA*IR is computed as using code in FA*IR codes. Proportion is implemented as statistical test 4.1.3 in Proportion paper.

http://demo.dataresponsibly.com/rankingfacts/nutrition_facts/

[K. Yang, J. Stoyanovich, A. Asudeh, B. Howe, HV Jagadish, G. Miklau; *SIGMOD 2018*]

Responsible data science

Because of its tremendous **power**, data science must be used **responsibly**

fairness

diversity

transparency

data protection