Responsible Data Science

Transparency in online ad delivery

Prof. George Wood

Center for Data Science
New York University

r/ai

NYU | Center for Data Science

This week's reading

Amit Datta*, Michael Carl Tschantz, and Anupam Datta

Automated Experiments on Ad Privacy Settings

A Tale of Opacity, Choice, and Discrimination

Google ads, black names and white names, racial discrimination, and click advertising.

BY LATANYA SWEENEY

Discrimination in Online Ad Delivery

Discrimination through Optimization:
How Facebook's Ad Delivery Can Lead to Biased Outcomes

MUHAMMAD ALI*, Northeastern University, USA
PIOTR SAPIEZYNSKI*, Northeastern University, USA
MIRANDA BOGEN, Upturn, USA
ALEKSANDRA KOROLOVA, University of Southern California, USA
ALAN MISLOVE, Northeastern University, USA
AARON RIEKE, Upturn, USA

r/ai

Instant Checkmate

Latanya Sweeney's experiment

Figure 4. Sample face images on google.com retrieved for searches "latanya" (a), "latisha" (b), "kristen" (c), and "jill" (d).

February 2013

Findings

Ads suggestive of criminal record, linking to Instant Checkmate, appear on google.com and reuters.com in response to searches "Latanya Sweeney", "Latanya Farrell", and "Latanya Locket"*

No Instant Checkmate ads when searching for "Kristen Haring", "Kristen Sparrow"*, and "Kristen Lindquist"*
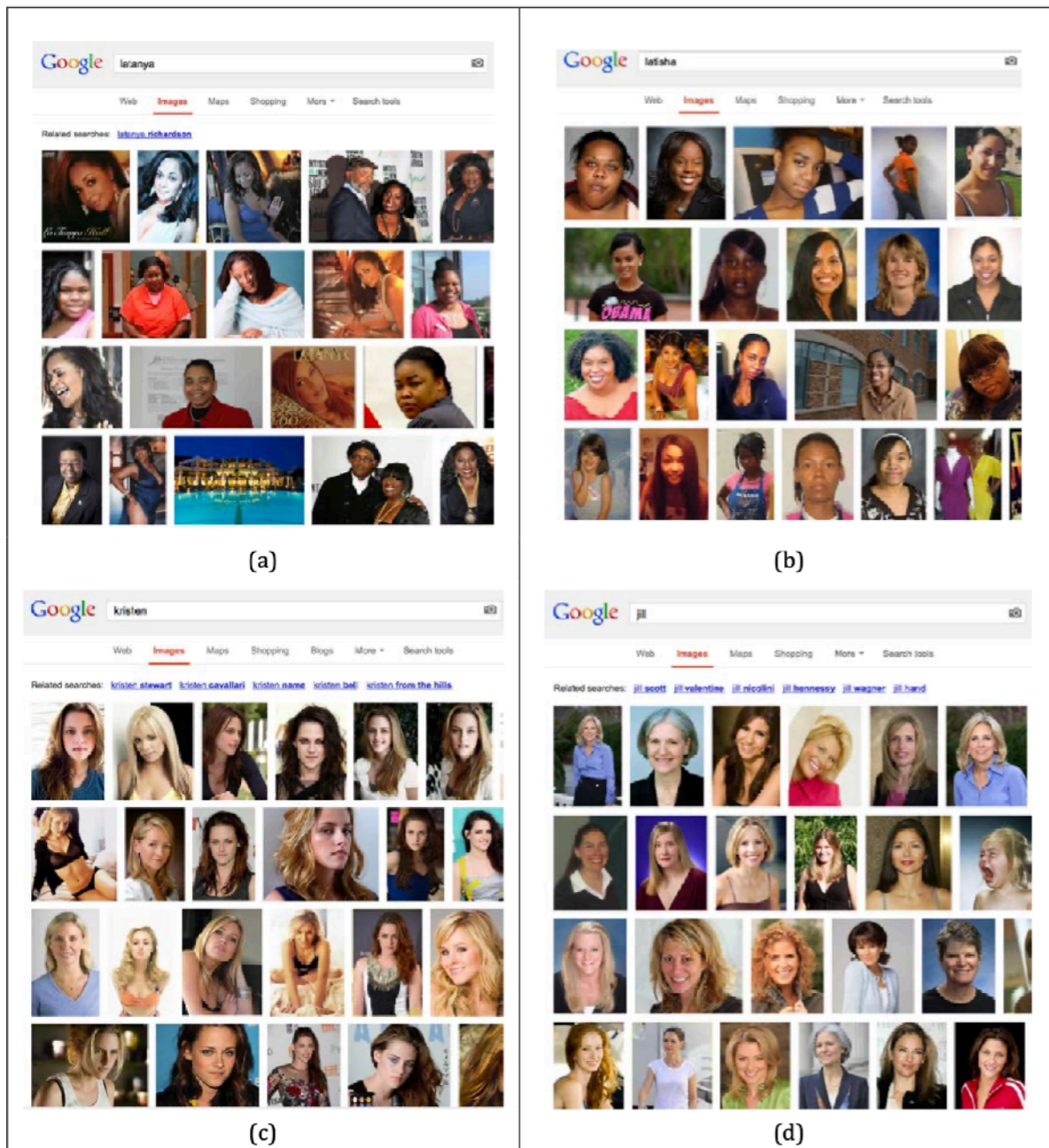
* Name associated with an actual arrest record

r/ai

Possible explanations

Conjectures

Does Instant Checkmate serve ads specifically for Black-identifying names?

Is Google AdSense explicitly biased in this way?

Does Google AdSense learn racial bias from click-through rates?

February 2013

Response

Google:"AdWords does not conduct any racial profiling. …It is up to individual advertisers to decide which keywords they want to choose to trigger their ads."

"**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."

FALAAH ARIF KHAN

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

r/ai

Ad targeting online

Users browse the web, consume content, consume ads (i.e. view, click, purchase)

Publishers (e.g. The Guardian) host online content that often includes ads. They outsource ad placement to third-party **ad networks** (e.g. Google Ads)

Advertisers seek to place their ads on **publisher's** website (usually by bidding in ad auctions)

Ad networks track users across sites, collecting data. They connect **advertisers** and **publishers**.
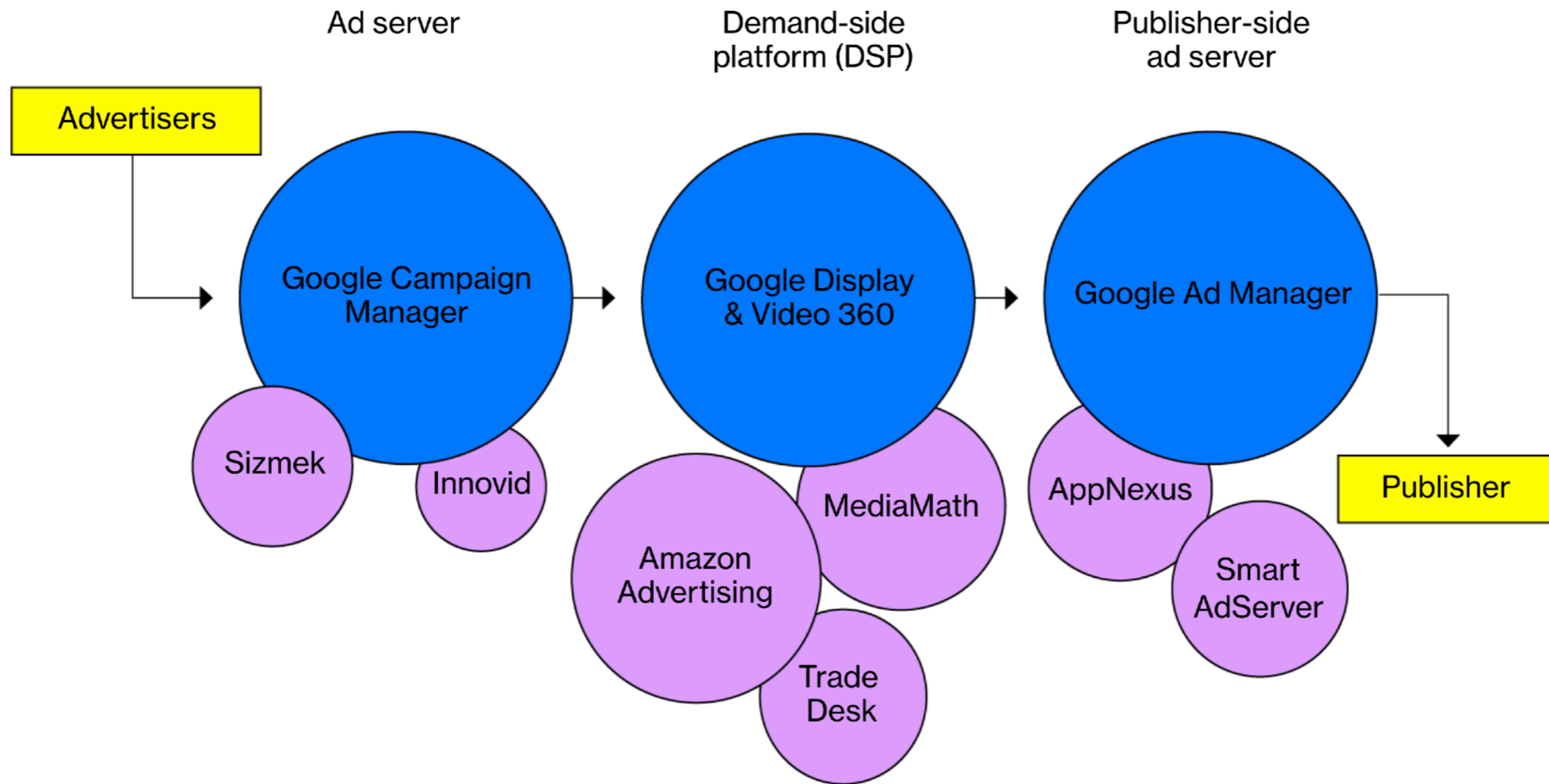
r/ai

Ad targeting online

https://www.bloomberg.com/news/features/2019-10-24/
how-google-s-ad-ecosystem-works

Google ad settings

Do users truly have transparency and choice, or are these placebo buttons?

Ad personalization

Google makes your ads more useful on Google services (such as Search or YouTube).

Ad personalization is ON

Control ad personalization on other websites & apps that use Google ad services.

r/ai

Google ad settings

Do users truly have transparency and choice, or are these placebo buttons?

Ad categories on YouTube BETA

These ad categories are not personalized, but you may see related ads based on your current searches or the topic of a video you're watching. You can choose to see fewer ads from these categories on YouTube.
Learn more about these ad categories

🍾	Alcohol	See fewer
⚅	Gambling	See fewer
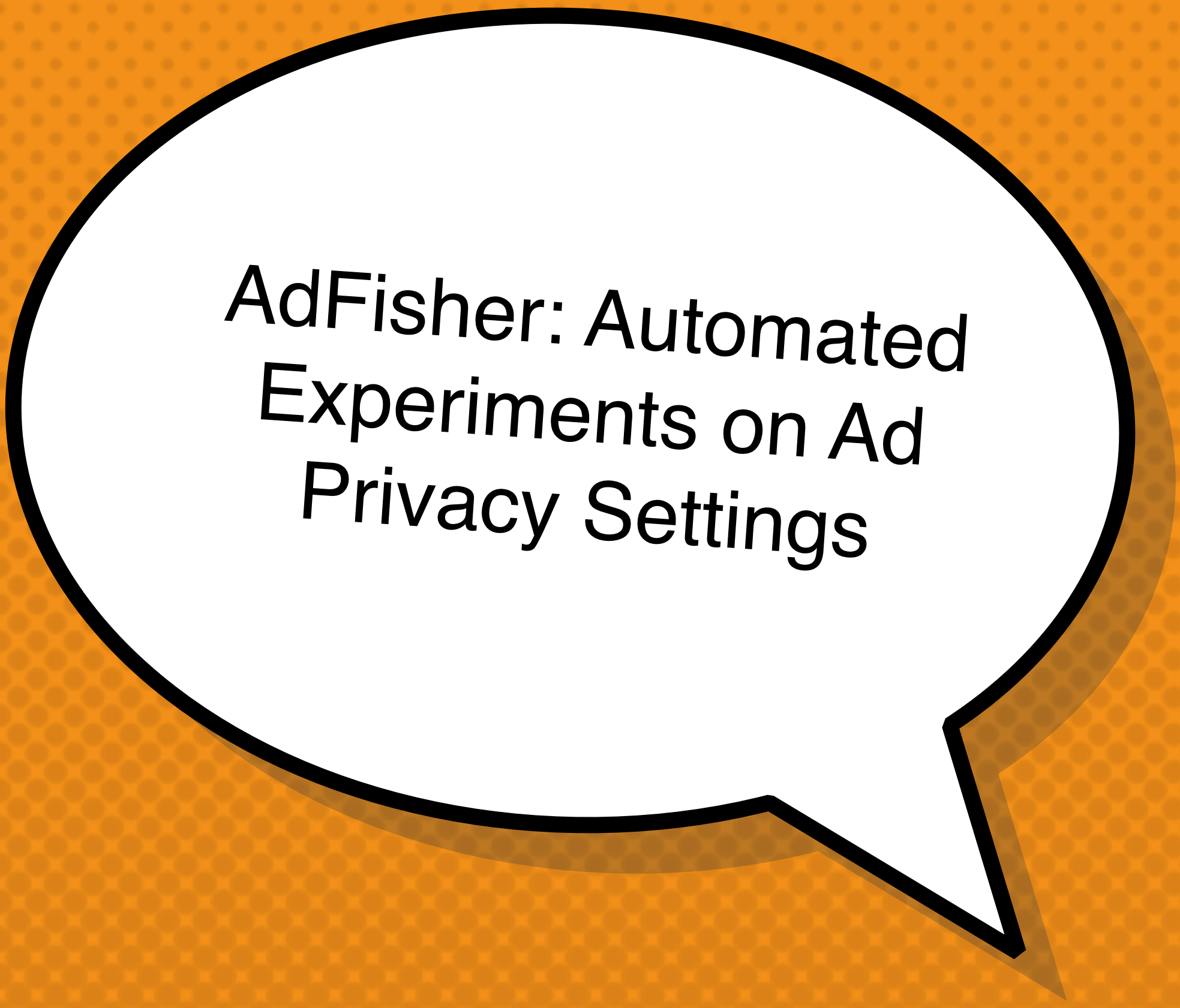
Google is one of many ad networks that personalizes ads based on your activity online. Go to **AdChoices** to control ads from other ad networks.

You can install a browser plugin to maintain your preference to opt out of personalized ads from Google, even if you've cleared your cookies. **Learn how to save settings for the browser**

We protect your privacy. Content from Drive, Gmail, and Photos is never used for any ads purposes.

Your data and ads   **Privacy Policy**

r/ai

AdFisher

theguardian

Samuel Gibbs

Wednesday 8 July 2015 11.29 BST

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

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

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

r/ai

AdFisher

Experiment

How do user behaviors, ads, and settings interact?

Individual data use transparency: ad network must disclose which user information is used when determining which ads to serve

[Datta, Tschantz, Datta, *PETS* 2015]
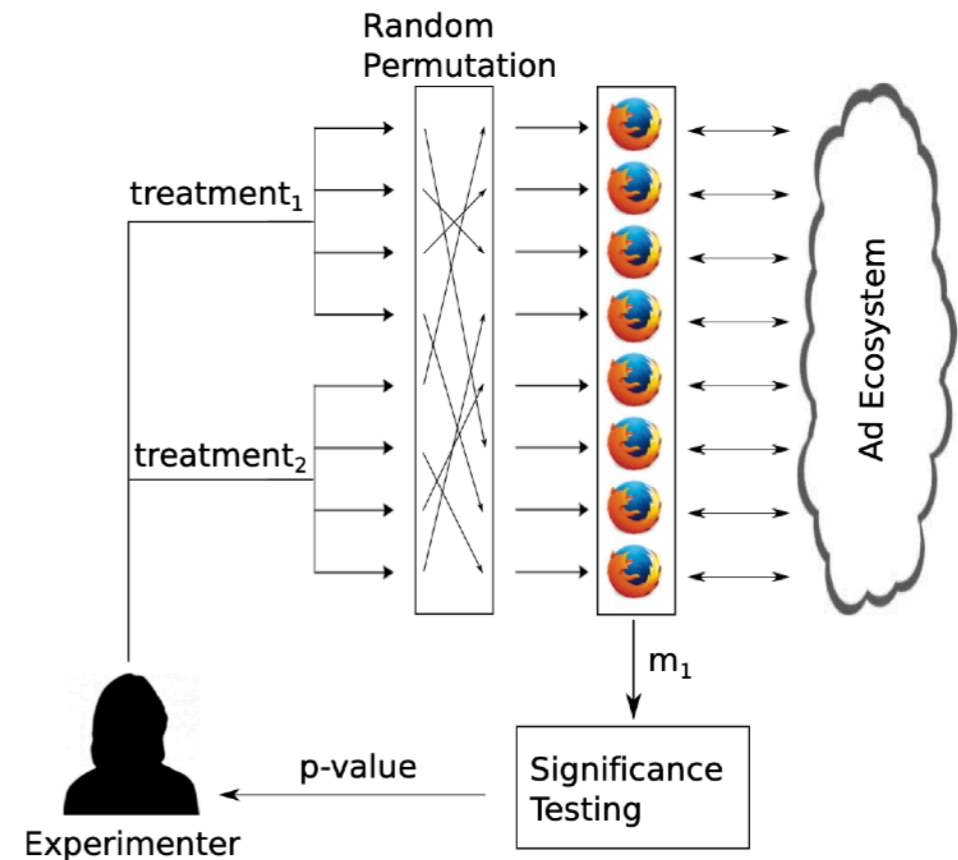
Figure 2: Experimental setup to carry out significance testing on eight browser agents comparing the effects of two treatments. Each agent is randomly assigned a treatment which specifies what actions to perform on the web. After these actions are complete, they collect measurements which are used for significance testing.

AdFisher: transparency tests

▶ Transparency: User can view data used for ad selection

▶ Causal test: Find attribute that changes ads but not settings

▶ Experiment 1: 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 ads served: rehab vs. stocks & driving jobs

[Datta, Tschantz, Datta, *PETS* 2015]

AdFisher: discrimination tests

▸ Non-Discrimination: Users differing only in protected attributes are treated similarly.

▸ Causal test: Does a protected attribute change ads?

▸ Experiment 2: 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: In one experiment, males were shown ads for higher-paying jobs far more often than females

[Datta, Tschantz, Datta, *PETS* 2015]

r/ai

Title	URL	Coefficient	appears in agents		total appearances	
			female	male	female	male
Top ads for identifying the simulated male group						
Truck Driving Jobs	www.bestpayingdriverjobs.com	0.492	0	15	0	33
$200k+ Jobs - Execs Only	careerchange.com	0.383	0	15	0	48
Aircraft Technician Program	pia.edu	0.292	0	6	0	14
Paid CDL Training	pamtransport.greatcdltraining.com	0.235	0	5	0	13
Unique Bridal Necklaces	margaretelizabeth.com/Bridal	0.234	0	5	0	19
Top ads for identifying agents in the simulated female group						
Business Loans for Women	topsbaloans.com	−0.334	13	1	70	1
Post Your Classified Ad	indeed.com/Post-Jobs	−0.267	20	16	56	24
American Laser Skincare	americanlaser.com	−0.243	8	5	14	8
Dedicated Drivers Needed	warrentransport.com	−0.224	3	0	14	0
Earn Your Nursing Degree	nursing-colleges.courseadvisor.com	−0.219	11	3	31	10

Table 5: Top URL+titles for the gender and jobs experiment (July).

[Datta, Tschantz, Datta, *PETS* 2015]

Follow-up: How is targeting done?

- On gender directly

- On a proxy of gender, i.e., on a known correlate of gender because it is a correlate

- On a known correlate of gender, but not because it is a correlate

- On an unknown correlate of gender
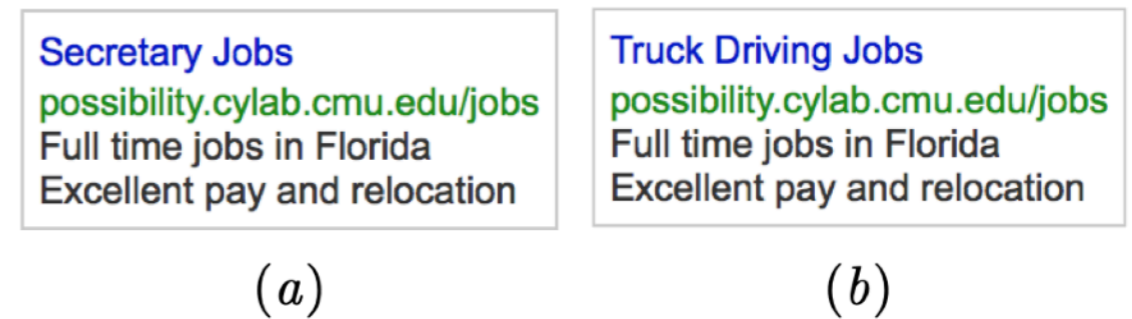
Figure 1: Ads approved by Google in 2015. The ad in the left (right) column was targeted to women (men).

"This finding demonstrates that an advertiser with discriminatory intentions can use the AdWords platform to serve employment related ads disparately on gender."

[A. Datta, A. Datta, J. Makagon, D. Mulligan, M. Tschantz; *FAT* 2018*]

AdFisher: Who is responsible?

Finding

Secretary Jobs	Truck Driving Jobs
possibility.cylab.cmu.edu/jobs	possibility.cylab.cmu.edu/jobs
Full time jobs in Florida	Full time jobs in Florida
Excellent pay and relocation	Excellent pay and relocation
(a)	(b)

Figure 1: Ads approved by Google in 2015. The ad in the left (right) column was targeted to women (men).

Conjectures

Is Google explicitly programming the system to show the ad less often to women?

Is the advertiser targeting the ad through explicit use of demographic categories or selection of proxies, and Google respecting these targeting criteria?

Are other advertisers: outbidding our advertiser when targeting to women?

Are male and female users behaving differently in response to ads?

[Datta, Datta, Makagon, Mulligan & Tschantz, 2018]

r/ai

Discrimination through optimization

▸ Follow-up work on AdFisher (Google ads, gender-based discrimination for the purposes of employment) ascertained that it was possible to target on gender for job ads

▸ Platforms have since taken steps to address such blatant violations

"… Facebook currently has several policies in place to avoid discrimination for certain types of ads. Facebook also recently built tools to automatically detect ads offering housing, employment, and credit, and pledged to prevent the use of certain targeting categories with those ads. Additionally, Facebook relies on advertisers to self-certify that they are not in violation of Facebook's advertising policy prohibitions against discriminatory practices. More recently, in order to settle multiple lawsuits stemming from these reports, Facebook stated that they will soon no longer allow age, gender, or ZIP code-based targeting for housing, employment or credit ads, and that they would also block other detailed targeting attributes that are "describing or appearing to relate to protected classes".

▸ Yet, the question remains: Does the ad delivery platform itself embed discriminatory outcomes?

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

r/ai

Discrimination through optimization

Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes

Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, Aaron Rieke

The enormous financial success of online advertising platforms is partially due to the precise targeting features they offer. Although researchers and journalists have found many ways that advertisers can target---or exclude---particular groups of users seeing their ads, comparatively little attention has been paid to the implications of the platform's ad delivery process, comprised of the platform's choices about which users see which ads.

It has been hypothesized that this process can "skew" ad delivery in ways that the advertisers do not intend, making some users less likely than others to see particular ads based on their demographic characteristics. In this paper, we demonstrate that such skewed delivery occurs on Facebook, due to market and financial optimization effects as well as the platform's own predictions about the "relevance" of ads to different groups of users. We find that both the advertiser's budget and the content of the ad each significantly contribute to the skew of Facebook's ad delivery. Critically, we observe significant skew in delivery along gender and racial lines for "real" ads for employment and housing opportunities despite neutral targeting parameters.

r/ai

Discrimination through optimization

Key question: does the platform itself introduce demographic skew in ad delivery?

Conjectures

Users see relevant ads, maximizing the likelihood of engagement. Based on historical engagement data, delivery may be skewed in ways that an advertiser may not have intended.

Market effects and financial optimization can lead to skewed ad delivery. In a nutshell: some populations are more "valuable" and so advertising to them costs more. If an advertiser bids less, they won't get to the more "valuable" population.

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

r/ai

Discrimination through optimization

Part 1: ad creation

- ad contents

- audience selection

- bidding strategy

Part 2: ad delivery

For every opportunity to show a user an ad (e.g., **an ad slot** is available as the user is browsing the service), the ad platform will run an **ad auction** to determine, from among all of the ads that include the current user in the audience, which ad should be shown.

Figure 1: Each ad has five elements that the advertiser can control: (1) the ad text, entered manually by the advertiser, (2) the images and/or videos, (3) the domain, pulled automatically from the HTML `meta` property `og:site_name` of the destination URL, (4) the title, pulled automatically from the HTML `meta` property `og:title` of the destination URL, and (5) the description from `meta` property `og:description` of the destination URL. The title and description can be manually customized by the advertiser if they wish.

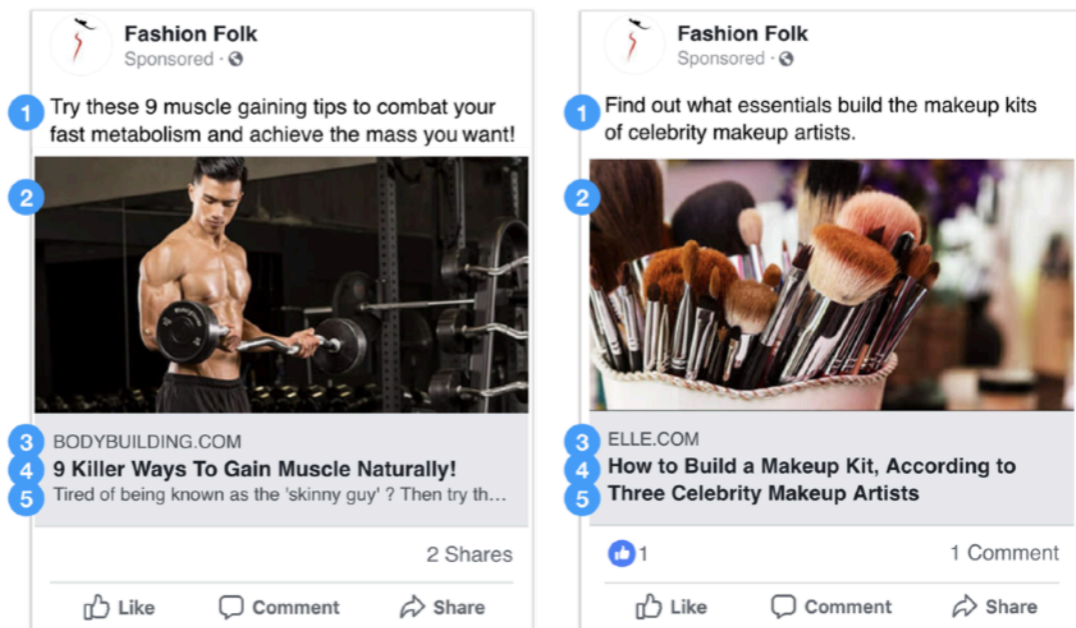[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

Discrimination through optimization

Part 1: ad creation

- ad contents

- audience selection

- bidding strategy

Part 2: ad delivery

For every opportunity to show a user an ad (e.g., **an ad slot** is available as the user is browsing the service), the ad platform will run an **ad auction** to determine, from among all of the ads that include the current user in the audience, which ad should be shown.

When Facebook has ad slots available, it runs an ad auction among the active advertisements bidding for that user. However, **the auction does not just use the bids placed by the advertisers**; Facebook says:

*"The ad that wins an auction and gets shown is the one with the highest **total value**. Total value isn't how much an advertiser is willing to pay us to show their ad. It's combination of 3 major factors: (1) Bid, (2) Estimated action rates, and (3) Ad quality and relevance."*

*"During ad set creation, you chose a target audience ... and an optimization event ... **We show your ad to people in that target audience who are likely to get you that optimization event.**"*

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

r/ai

Discrimination through optimization

Facebook ad delivery results can be skewed in ways that advertisers do not intend

Skew can arise due to:

- ‣ Financial optimization effects

- ‣ The ad delivery platform's predictions about the relevance of its ads to different user categories
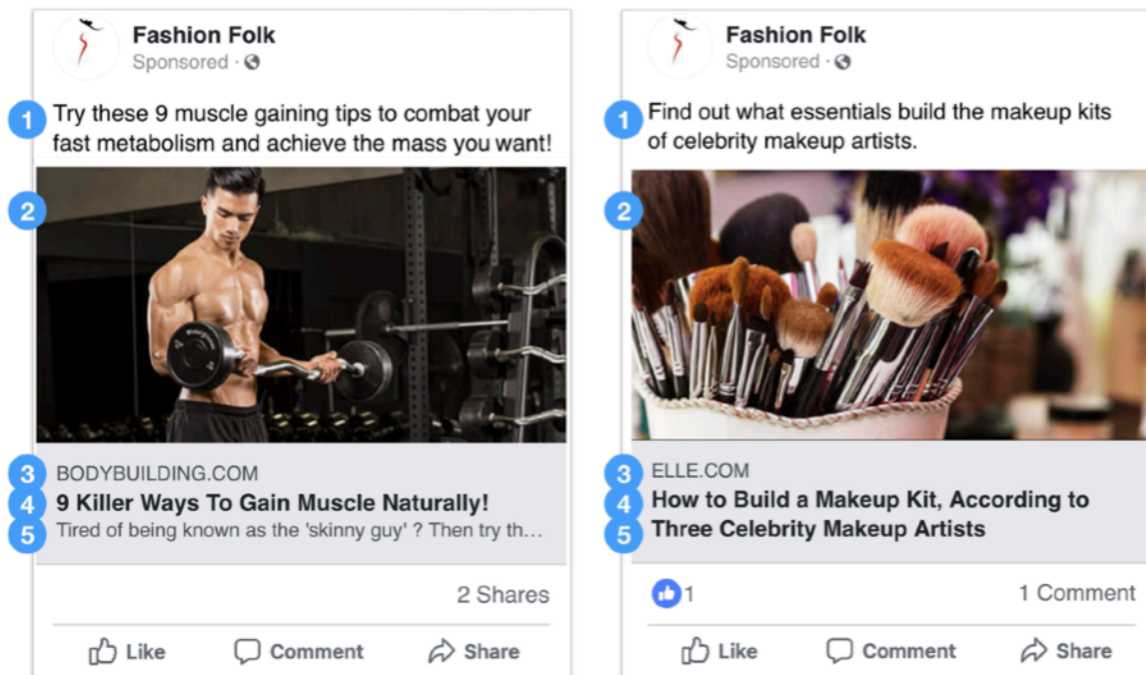
What contributes to the skew?

- ‣ Ad content; both text and images, which are likely automatically analyzed by Facebook

- ‣ Advertiser budget

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

r/ai

Discrimination through optimization
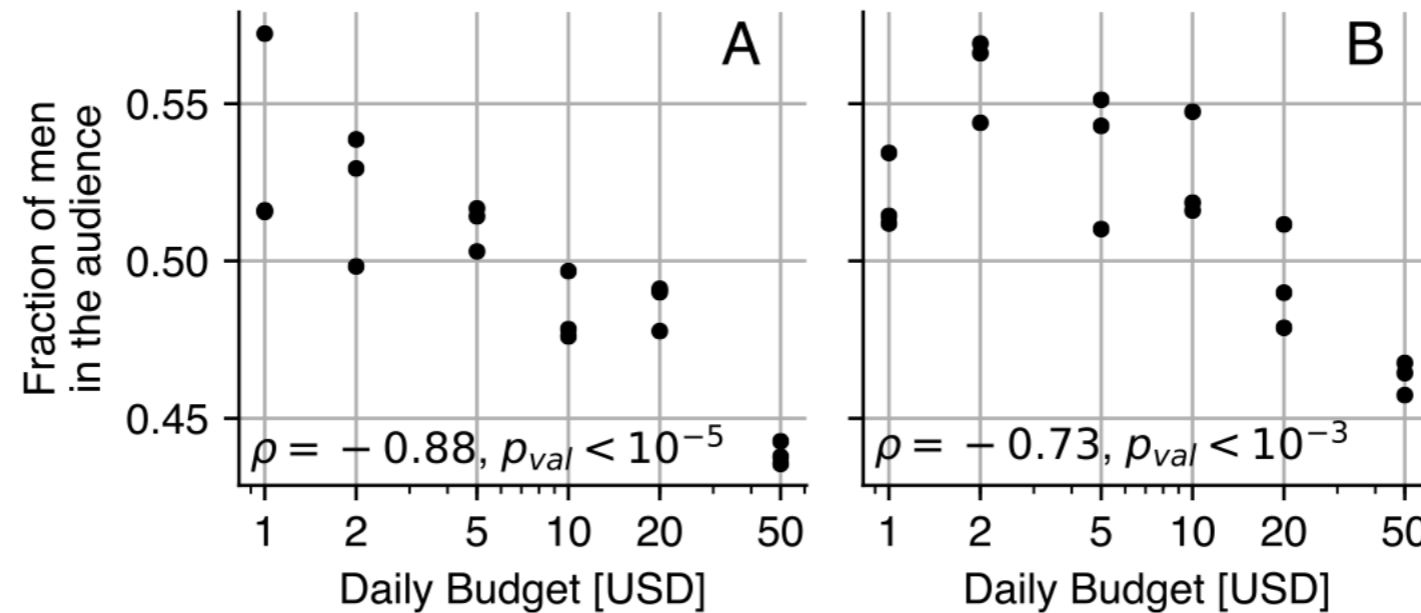
Figure 2: Gender distributions of the audience depend on the daily budget of an ad, with higher budgets leading to a higher fraction of women. The left graph shows an experiment where we target all users located in the U.S.; the right graph shows an experiment where we target our random phone number custom audiences.

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

Discrimination through optimization

Figure 1: Each ad has five elements that the advertiser can control: (1) the ad text, entered manually by the advertiser, (2) the images and/or videos, (3) the domain, pulled automatically from the HTML meta property og:site_name of the destination URL, (4) the title, pulled automatically from the HTML meta property og:title of the destination URL, and (5) the description from meta property og:description of the destination URL. The title and description can be manually customized by the advertiser if they wish.

Same bidding strategy for bodybuilding and cosmetics, without explicitly mentioning gender

Strong gender skew in delivery: bodybuilding delivered to over 75% men on average, cosmetics delivered to over 90% women on average

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

Discrimination through optimization
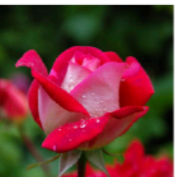
Which component of the ad creative impacts delivery most?
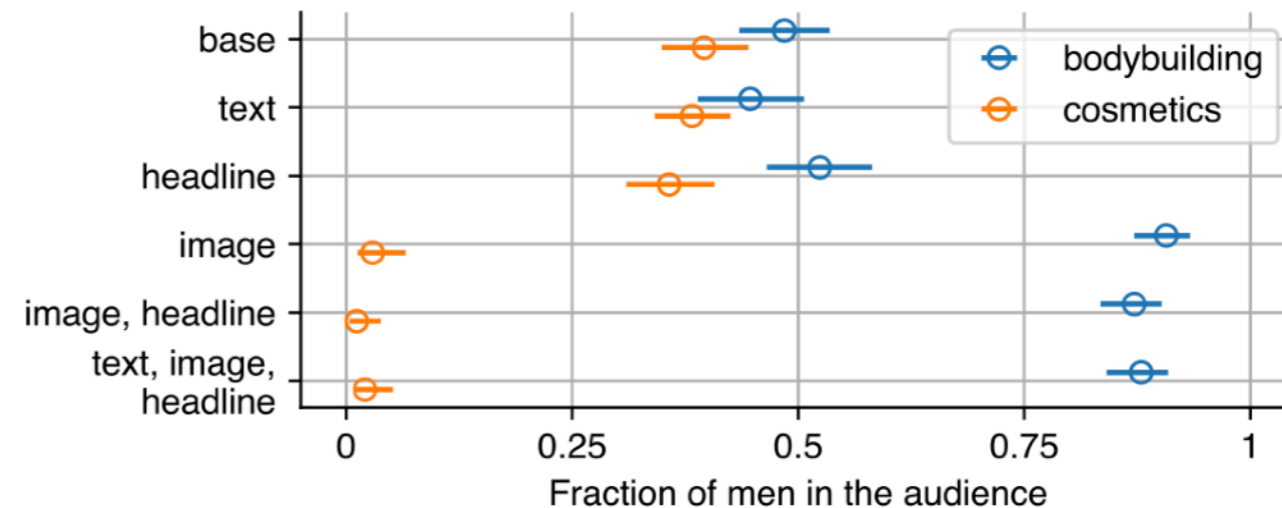
Figure 3: "Base" ad contains a link to a page about either bodybuilding or cosmetics, a blank image, no text, or headline. There is a small difference in the fraction of male users for the base ads, and adding the "text" only decreases it. Setting the "headline" sets the two ads apart but the audience of each is still not significantly different than that of the base version. Finally, setting the ad "image" causes drastic changes: the bodybuilding ad is shown to a 91% male audience, the cosmetics ad is shown to a 5% male audience, despite the same target audience.

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

Discrimination through optimization

Transparent images are still targeted

This strongly suggests that Facebook uses **an automated image classification mechanism** to steer different ads towards different subsets of the user population

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

Discrimination through optimization
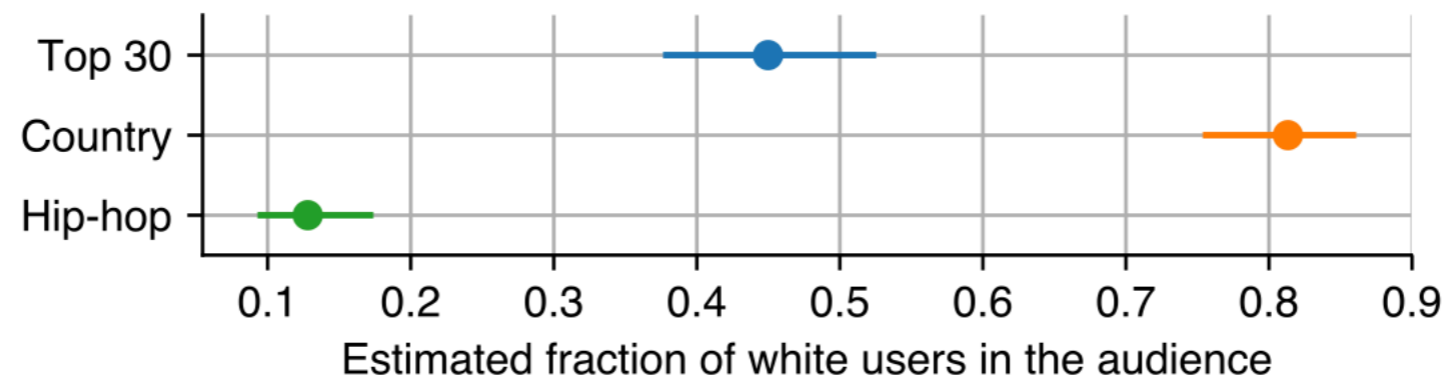
Figure 7: We run three campaigns about the best selling albums. *Top 30* is neutral, targeting all. *Country* implicitly targets white users, and *Hip-hop* implicitly targets Black users. Facebook classification picks up on the implicit targeting and shows it to the audience we would expect.

"We hold targeting parameters fixed, run ads that are stereotypically of interest to different races. We find that **Facebook ad delivery follows the stereotypical distribution**, despite all ads being targeted in the same manner and using the same bidding strategy."
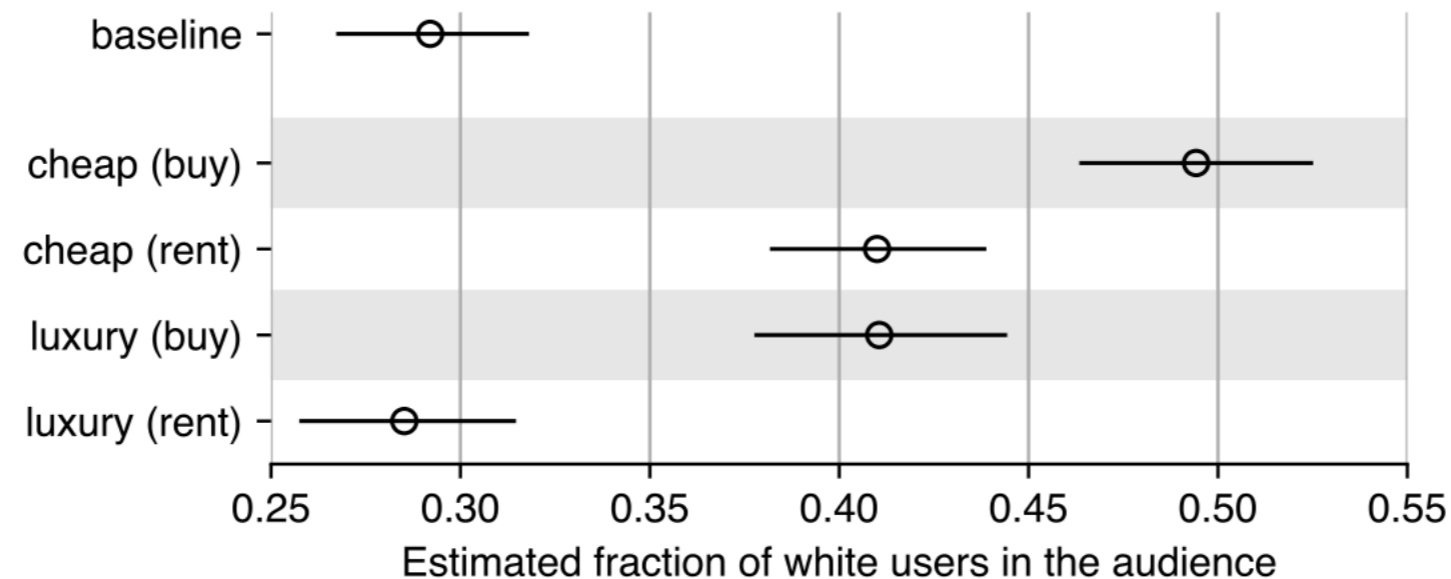
[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

Discrimination through optimization

Figure 9: Results for housing ads, showing a breakdown in the ad delivery audience by race. Despite being targeted in the same manner, using the same bidding strategy, and being run at the same time, we observe significant skew in the makeup of the audience to whom the ad is delivered (ranging from estimated 27% white users for luxury rental ads to 49% for cheap house purchase ads).

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

Discrimination through optimization

Key question: does the platform itself introduce demographic skew in ad delivery?

Findings

Skew can arise due to financial optimization effects and the ad delivery platform's predictions about the relevance of its ads to different user categories

Ad content - text and images - and advertiser budget both may contribute to the skew.

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

r/ai

Discrimination through optimization

Key question: does the platform itself introduce demographic skew in ad delivery?
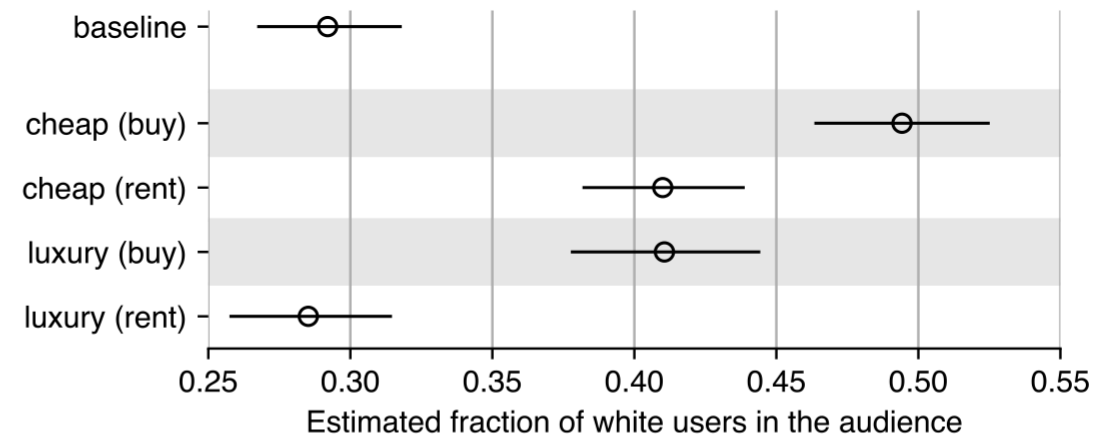
Figure 9: Results for housing ads, showing a breakdown in the ad delivery audience by race. Despite being targeted in the same manner, using the same bidding strategy, and being run at the same time, we observe significant skew in the makeup of the audience to whom the ad is delivered (ranging from estimated 27% white users for luxury rental ads to 49% for cheap house purchase ads).

Findings

Skew was observed along racial lines, in ads for housing opportunities

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

When is "skew" in fact discrimination?

SUMMARY OF SETTLEMENTS BETWEEN CIVIL RIGHTS ADVOCATES

AND FACEBOOK

Housing, Employment, and Credit Advertising Reforms

In the settlements, Facebook will undertake far-reaching changes and steps that will prevent discrimination in housing, employment, and credit advertising on Facebook, Instagram, and Messenger. These changes demonstrate real progress.

- **Facebook will establish a separate advertising portal for creating housing, employment, and credit ("HEC") ads on Facebook, Instagram, and Messenger that will have limited targeting options, to prevent discrimination.**

- **The following rules will apply to creating HEC ads.**

  o *Gender, age, and multicultural affinity targeting options will not be available when creating Facebook ads.*

  o *HEC ads must have a minimum geographic radius of 15 miles from a specific address or from the center of a city. Targeting by zip code will not be permitted.*

March 19, 2019

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

r/ai

Facebook ads and HEC

SUMMARY OF SETTLEMENTS BETWEEN CIVIL RIGHTS ADVOCATES

AND FACEBOOK

Housing, Employment, and Credit Advertising Reforms

o *HEC ads will not have targeting options that describe or appear to be related to personal characteristics or classes protected under anti-discrimination laws. This means that targeting options that may relate to race, color, national origin, ethnicity, gender, age, religion, family status, disability, and sexual orientation, among other protected characteristics or classes, will not be permitted on the HEC portal.*

o *Facebook's "Lookalike Audience" tool, which helps advertisers identify Facebook users who are similar to advertisers' current customers or marketing lists, will no longer consider gender, age, religious views, zip codes, Facebook Group membership, or other similar categories when creating customized audiences for HEC ads.*

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

Facebook ads and HEC

SUMMARY OF SETTLEMENTS BETWEEN CIVIL RIGHTS ADVOCATES

AND FACEBOOK

Housing, Employment, and Credit Advertising Reforms

o *Advertisers will be asked to create their HEC ads in the HEC portal, and if Facebook detects that an advertiser has tried to create an HEC ad outside of the HEC portal, Facebook will block and re-route the advertiser to the HEC portal with limited options.*

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

r/ai

Legal implications

HUD reportedly also investigating Google and Twitter in housing discrimination probe

By Adi Robertson | @thedextriarchy | Mar 28, 2019, 3:52pm EDT

Fair Housing Act, also called **Title VIII of the Civil Rights Act of 1968**, U.S. federal legislation that protects individuals and families from discrimination in the sale, rental, financing, or **advertising** of housing. The Fair Housing Act, as amended in 1988, prohibits discrimination on the basis of **race**, **color**, **religion**, **sex**, **disability**, **family status**, and **national origin**.

Facebook has been charged with housing discrimination by the US government

'Facebook is discriminating against people based upon who they are and where they live,'
says HUD secretary

By Russell Brandom | Mar 28, 2019, 7:51am EDT

This is the first federal discrimination lawsuit to deal with **racial bias in targeted advertising**, a milestone that lawyers at HUD said was overdue. "Even as we confront new technologies, the fair housing laws enacted over half a century ago remain clear—discrimination in housing-related advertising is against the law," said HUD General Counsel Paul Compton. "**Just because a process to deliver advertising is opaque and complex doesn't mean that it's exempts Facebook and others from our scrutiny and the law of the land**."

[Datta, Datta, Makagon, Mulligan & Tschantz, 2018]

r/ai

Facebook ads and The Fair Housing Act

THE VERGE

Facebook has been charged with housing discrimination by the US government

'Facebook is discriminating against people based upon who they are and where they live,' says HUD secretary

By Russell Brandom | Mar 28, 2019, 7:51am EDT

Facebook has struggled to effectively address the possibility of discriminatory ad targeting. The company pledged to step up anti-discrimination enforcement in the wake of *ProPublica*'s reporting, but a follow-up report in 2017 found the same problems persisted nearly a year later.

"WE'RE DISAPPOINTED BY TODAY'S DEVELOPMENTS," FACEBOOK SAYS

According to the HUD complaint, many of the options for targeting or excluding audiences are shockingly direct, including a map tool that explicitly echoes redlining practices. "[Facebook] has provided a toggle button that enables advertisers to exclude men or women from seeing an ad, a search-box to exclude people who do not speak a specific language from seeing an ad, and a map tool to exclude people who live in a specified area from seeing an ad by drawing a red line around that area," the complaint reads.

https://www.theverge.com/2019/3/28/18285178/facebook-hud-lawsuit-fair-housing-discrimination

[Ali, Sapiezynski, Bogen, Korolova, Mislove & Rieke, 2019]

r/ai

The socio-legal landscape

Related concern:
Are ads commercial free speech?

▸ The First Amendment of the U.S. Constitution protects advertising, but the U.S. Supreme Court set out a test for assessing restrictions on commercial speech, which begins by determining whether the speech is misleading

▸ Are online ads suggesting the existence of an arrest record misleading if no one by that name has an arrest record?

▸ Assume the ads are free speech: what happens when these ads appear more often for one racial group than another? Not everyone is being equally affected. Is that free speech or racial discrimination?

[Sweeney, 2013]

r/ai

Tracking and consent

The New York Times

To Be Tracked or Not? Apple Is Now Giving Us the Choice.
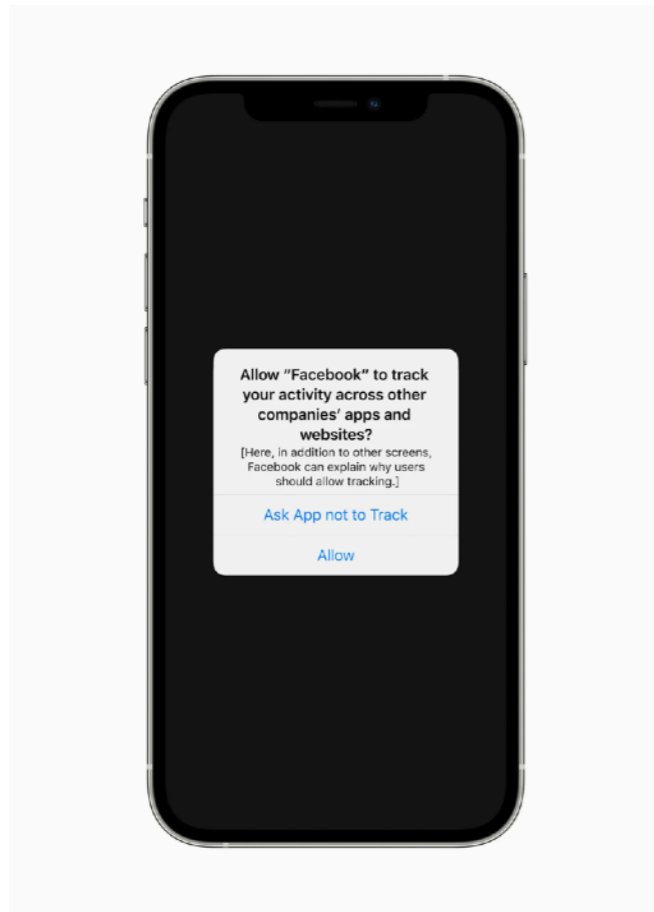
By **Brian X. Chen**

April 26, 2021 Updated 12:40 p.m. ET

If we had a choice, would any of us want to be tracked online for the sake of seeing more relevant digital ads?
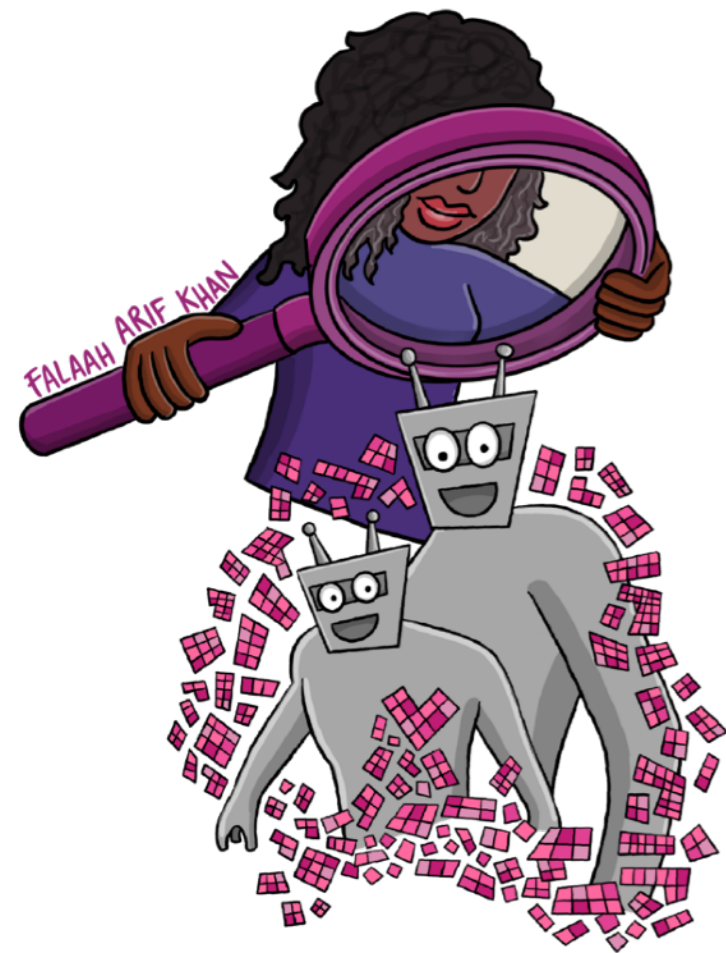
We are about to find out.

On Monday, Apple plans to release iOS 14.5, one of its most anticipated software updates for iPhones and iPads in years. It includes a new privacy tool, called App Tracking Transparency, which could give us more control over how our data is shared.

Here's how it works: When an app wants to follow our activities to share information with third parties such as advertisers, a window will show up on our Apple device to ask for our permission to do so. If we say no, the app must stop monitoring and sharing our data.

A pop-up window may sound like a minor design tweak, but it has thrown the online advertising industry into upheaval. Most notably, Facebook has gone on the warpath. Last year, the social network created a website and took out full-page ads in newspapers denouncing Apple's privacy feature as harmful to small businesses.

Allow "Facebook" to track your activity across other companies' apps and websites?

[Here, in addition to other screens, Facebook can explain why users should allow tracking.]
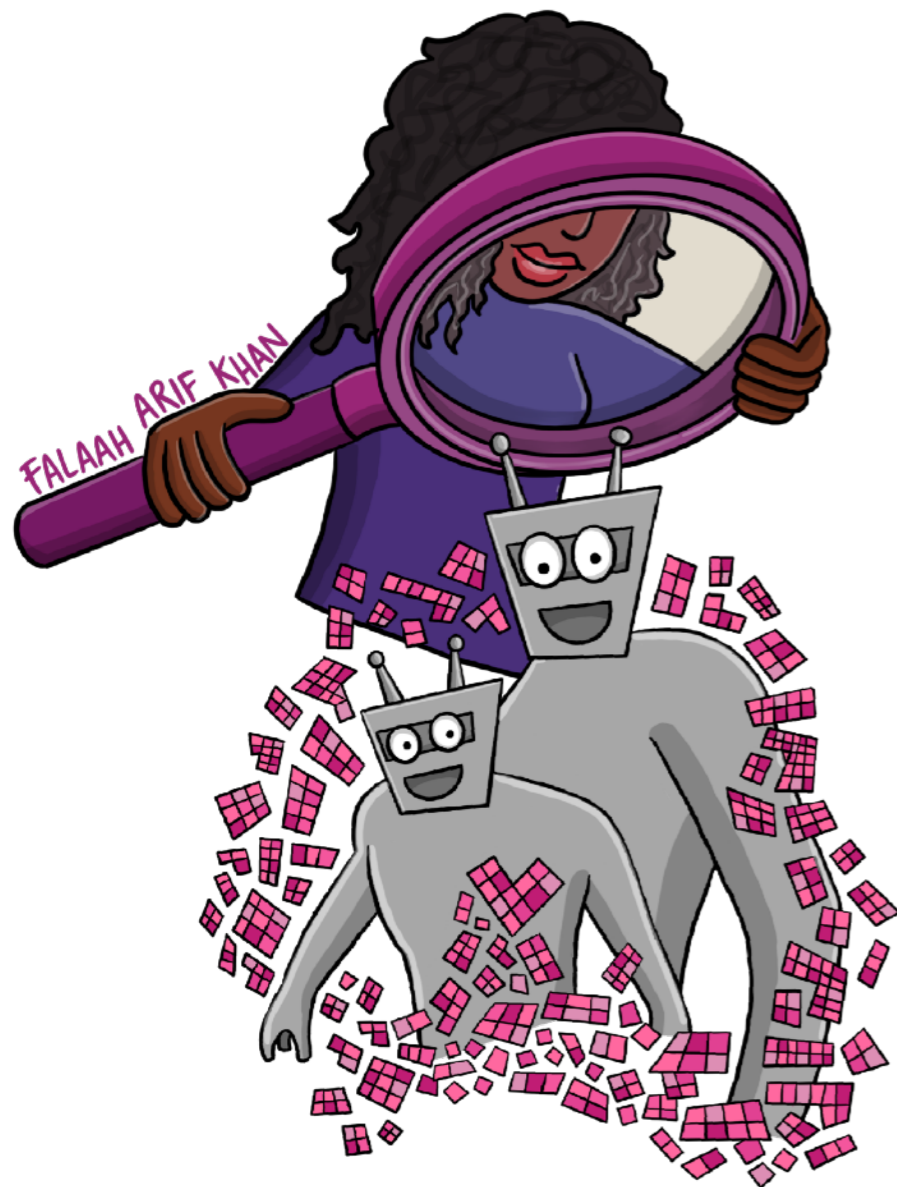
Ask App not to Track

Allow

r/ai

transparency, interpretability, explainability, intelligibility

responsible AI

agency, responsibility

r/ai

What are we explaining?

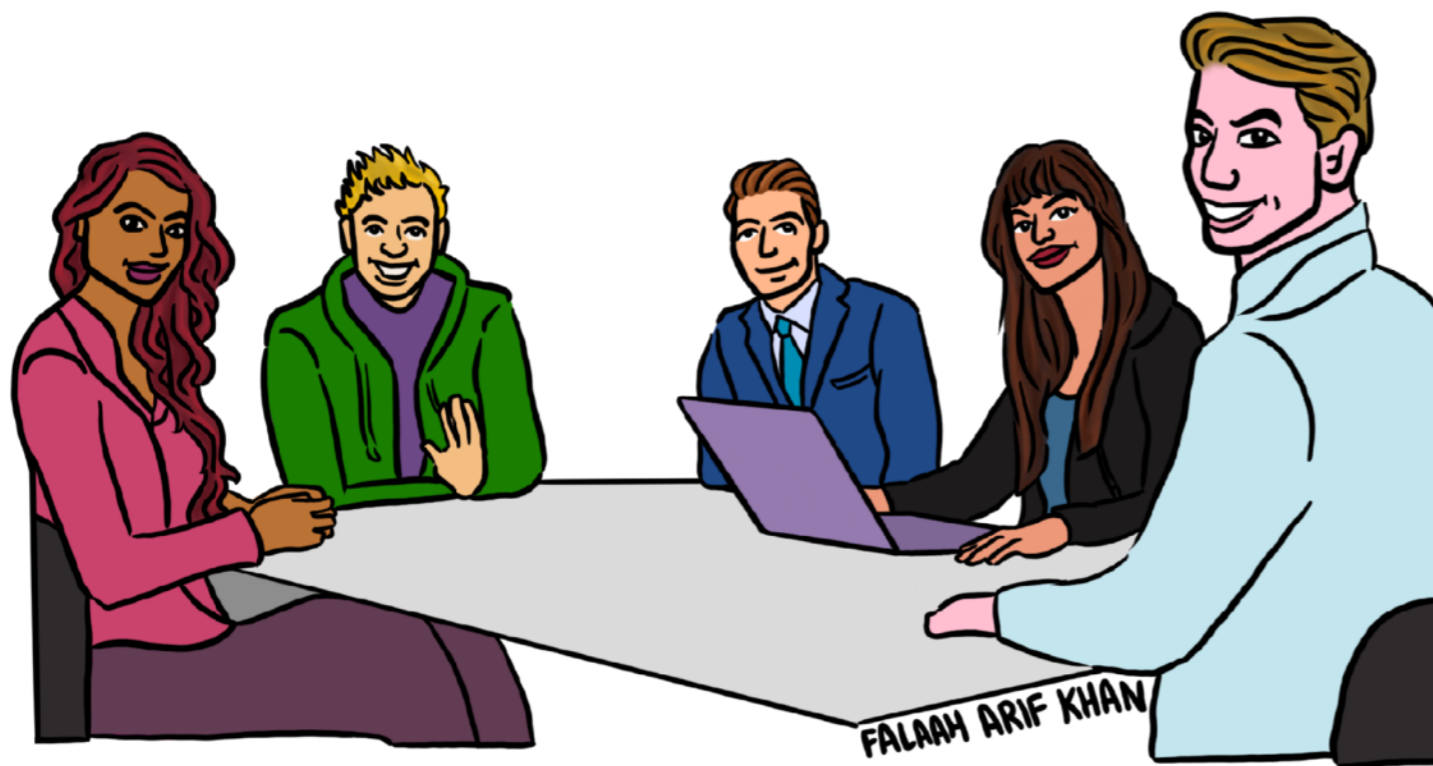How does a system work?

How **well** does a system work?

What does a system do?

Why was I ___ (mis-diagnosed / not offered a discount / denied credit) ?

Are a system's decisions discriminatory?

Are a system's decisions illegal?