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

BY LATANYA SWEENEY

Discrimination in Online Ad Delivery

DO ONLINE ADS suggestive of arrest records appear more often with searches of black-sounding names than white-sounding names? What is a black-sounding name or white-sounding name, anyway? How do you design technology to reason about societal consequences like structural racism? Let's take a scientific dive into online ad delivery to find answers.

"Have you ever been arrested?" Imagine this question appearing whenever someone enters your name in a search engine. Perhaps you are in competition for an award or a new job, or maybe you are in a position of trust, such as a professor or a volunteer. Perhaps you are dating or engaged in any one of hundreds of circumstances for which someone wants to learn more about you online. Appearing alongside your accomplishments is an advertisement implying you may have a criminal record, whether you actually have one or not. Worse, the ads may not appear for your competitors.

Employers frequently ask whether applicants have ever been arrested or charged with a crime, but if an employer disqualifies a job applicant based solely upon information indicating an arrest record, the company may face legal consequences. The U.S. Equal Employment Opportunity Commission (EEOC) is the federal agency charged with enforcing Title VII of the Civil Rights Act of 1964, a law that applies to most employers, prohibiting employment discrimination based on race, color, religion, sex, or national origin, and extended to those having criminal records.[5,11] Title VII does not prohibit employers from obtaining criminal background information, but a blanket policy of excluding applicants based solely upon information indicating an arrest record can result in a charge of discrimination.

To make a determination, the EEOC uses an adverse impact test that measures whether certain practices, intentional or not, have a disproportionate effect on a group of people whose defining characteristics are covered by Title VII. To decide, you calculate the percentage of people affected in each group and then divide the smaller value by the larger to get the ratio and compare the result to 80. If the ratio is less than 80, then the EEOC considers the effect disproportionate and may hold the employer responsible for discrimination.[6]

What about online ads suggesting someone with your name has an arrest record? Title VII only applies if you have an arrest record and can prove the employer inappropriately used the ads.

Are the ads commercial free speech—a constitutional right to display the ad associated with your name? 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.[3] Are online ads suggesting the existence of an arrest record misleading if no one by that name has an arrest record?

Figure 1. Ads from a Google search of three different names beginning with first name "Latanya."

(a)

(b)

(c)

(d)

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 by free speech. Is that free speech or racial discrimination?

Racism, as defined by the U.S. Commission on Civil Rights, is "any attitude, action, or institutional structure which subordinates a person or group because of their color."[16] *Racial discrimination* results when a person or group of people is treated differently based on their racial origins, according to the Panel on Methods for Assessing Discrimination of the National Research Council.[12] Power is a necessary precondition, for it depends on the ability to give or withhold benefits, facilities, services, and opportunities from someone who should be entitled to them and is denied on the basis of race. *Institutional or structural racism*, as defined in *The Social Work Dictionary*, is a system of procedures/patterns whose effect is to foster discriminatory outcomes or give preferences to members of one group over another.[1]

These considerations frame the relevant socio-legal landscape. Now we turn to whether online ads suggestive of arrest records appear more often for one racial group than another among a sample of racially associated names, and if so, how technology can solve the problem.

The Pattern

What is the suspected pattern of ad delivery? Here is an overview using real-world examples.

Earlier this year, a Google search for *Latanya Farrell*, *Latanya Sweeney*, and *Latanya Lockett* yielded ads and criminal reports like those shown in Figure 1. The ads appeared on Google.com (Figure 1a, 1c) and on a news website, Reuters.com, to which Google supplies ads (Figure 1c), All the ads in question linked to instantcheckmate.com (Figure 1b, 1d). The first ad implied *Latanya Farrell* might have been arrested. Was she? Clicking on the link and paying the requisite fee revealed the company had no arrest record for her or *Latanya Sweeney*, but there is a record for *Latanya Lockett*.

In comparison, searches for *Kristen Haring*, *Kristen Sparrow*, and *Kristen Lindquist* did not yield any instant-

checkmate.com ads, even though the company's database reported having records for all three names and arrest records for *Sparrow* and *Lindquist*.

Searches for *Jill Foley*, *Jill Schneider*, and *Jill James* displayed instantcheckmate.com ads with neutral copy; the word *arrest* did not appear in the ads even though arrest records for all three names appeared in the company's database. Figure 2 shows ads appearing on Google.com and Reuters.com and criminal reports from instantcheckmate.com for the first two names.

Finally, we considered a proxy for race associated with these names. Figure 3 shows racial distinction in Google image search results for *Latanya*, *Latisha*, *Kristen*, and *Jill*, respectively. The faces associated with *Latanya* and *Latisha* tend to be black, while white faces dominate the images of *Kristen* and *Jill*.

These handpicked examples describe the suspected pattern: ads suggesting arrest tend to appear with names associated with blacks, and neutral or no ads appear with names associated with whites, regardless of whether the company placing the ad has an arrest record associated with the name.

Google Adsense

Who generates the ad's text? Who decides when and where an ad will appear? What is the relationship among Google, a news website such as Reuters, and Instant Checkmate in the previous examples? An overview of Google AdSense, the program that delivered the ads, provides the answers.

In printed newspapers, everyone who reads the publication sees the same ad in the same space. Online ads can be tailored to the reader's search criteria, interests, geographical location, and so on. Any two readers (or even the same reader returning to the same website) might view different ads.

Google AdSense is the largest provider of dynamic online advertisements, placing ads for millions of sponsors on millions of websites.[9] In the first quarter of 2011, Google earned $2.43 billion through Google AdSense.[10] Several different advertising arrangements exist, but for simplicity this article describes only those features of Google AdSense specific to the Instant Checkmate ads in question.

Figure 2. Ad from a search of three different names beginning with the first name "Jill."

When a reader enters search criteria in an enrolled website, Google AdSense embeds into the Web page of results ads believed to be relevant to the search. Figures 1 and 2 show ads delivered by Google AdSense in response to various *firstname lastname* searches.

An advertiser provides Google with search criteria, copies of possible ads to deliver, and a bid to pay if a reader clicks the delivered ad. (For convenience, this article conflates Google AdSense with the related Google Adwords.) Google operates a real-time auction across bids for the same search criteria based on a "quality score" for each bid. A quality score includes many factors such as the past performance of the ad and characteristics of the company's website.[10] The ad having the highest quality score appears first, the second-highest second, and so on, and Google may elect not to show any ad if it considers the bid too low or if showing the ad exceeds a threshold (For example, a maximum account total for the advertiser). The Instant Checkmate ads in figures 1 and 2 often appeared first among ads, implying Instant Checkmate ads had the highest quality scores.

A website owner wanting to "host" online ads enrolls in AdSense and modifies the website to send a user's search criteria to Google and to display returning ads under a banner "Ads by Google" among search results. For example, Reuters.com hosts AdSense, and entering *Latanya Sweeney* in the search bar generated a new Web page with ads under the banner "Ads by Google" (Figure 1c).

There is no cost for displaying an ad, but if the user actually clicks on the ad, the advertiser pays the auction price. This may be as little as a few pennies, and the amount is split between Google and the host. Clicking the *Latanya Sweeney* ad on Reuters.com (Figure 1c) would cause Instant Checkmate to pay its auction amount to Google, and Google would split the amount with Reuters.

Search Criteria

What search criteria did Instant Checkmate specify? Will ads be delivered for made-up names? Ads displayed on Google.com allow users to learn why a specific ad appeared. Clicking the circled "i" in the ad banner (for example, Figure 1c) leads to a Web page explaining the ads. Doing so for ads in figures 1 and 2 reveals that the ads appeared because the search criteria matched the exact first- and last-name combination searched.

So, the search criteria must consist of both first and last names; and the names should belong to real people because a company presumably bids on records it sells.

The next steps describe the systematic construction of a list of racially associated first and last names for real people to use as search criteria. Neither Instant Checkmate nor Google are presumed to have used such a list.

Rather, the list provides a qualified sample of names to use in testing ad-delivery systems.

Black- and White-Identifying Names

Black-identifying and white-identifying first names occur with sufficiently higher frequency in one race than the other.

In 2003 Marianne Bertrand and Sendhil Mullainathan of the National Bureau of Economic Research (NBER) conducted an experiment in which they provided resumes to job posts that were virtually identical, except some of the resumes had black-identifying names and others had white-identifying names. Results showed white names received 50% more interviews.[2]

The study used names given to black and white babies in Massachusetts between 1974 and 1979, defining black-identifying and white-identifying names as those that have the highest ratio of frequency in one racial group to frequency in the other racial group.

In the popular book *Freakonomics*, Steven Levitt and Stephen Dubner report the top 20 whitest- and blackest-identifying girl and boy names. The list comes from earlier work by Levitt and Roland Fryer, which shows a pattern change in the way blacks named their children starting in the 1970s.[7] It was compiled from names given to black and white children recorded in California birth records from 1961–2000 (more than 16 million births).

To test ad delivery, I combined the lists from these prior studies and added two black female names, *Latanya* and *Latisha*. Table 1 lists the names used here, consisting of eight for each of the categories: white female, black female, white male, and black male from the Bertrand and Mullainathan study (first row in Table 1); and the first eight names for each category from the Fryer and Levitt work (second row in Table 1). Emily, a white female name, Ebony, a black female name, and Darnell, a black male name, appear in both rows. The third row includes the observation shown in Figure 3. Removing duplicates leaves a total of 63 distinct first names.

Full Names of Real People

Web searches provide a means of locating and harvesting a real person's first and last name (full name) by sampling

Table 1. Black-identifying names and white-identifying first names.

	White Female	Black Female	White Male	Black Male
(a)	Allison	Aisha	Brad	Darnell
	Anne	Ebony	Brendan	Hakim
	Carrie	Keisha	Geoffrey	Jermaine
	Emily	Kenya	Greg	Kareem
	Jill	Latonya	Brett	Jamal
	Laurie	Lakisha	Jay	Leroy
	Kristen	Latoya	Matthew	Rasheed
	Meredith	Tamika	Neil	Tremayne
(b)	Molly	Imani	Jake	DeShawn
	Amy	Ebony*	Connor	DeAndre
	Claire	Shanice	Tanner	Marquis
	Emily*	Aaliyah	Wyatt	Darnell*
	Katie	Precious	Cody	Terrell
	Madeline	Nia	Dustin	Malik
	Katelyn	Deja	Luke	Trevon
	Emma	Diamond	Jack	Tyrone
(c)		Latanya		
		Latisha		

names of professionals appearing on the Web; and sampling names of people active on social media sites and blogs (netizens).

Professionals often have their own Web pages that list positions and describe prior accomplishments. Several professions have degree designations (for example, Ph.D., M.D., J.D., or MBA) associated with people in that profession. A Google search for a first name and a degree designation can yield lists of people having that first name and degree.

The next step is to visit the Web page associated with each full name, and if an image is discernible, record whether the person appears black, white, or other.

Here are two examples from my test. A Google search for *Ebony PhD* revealed links for real people having *Ebony* as a first name—specifically, *Ebony Bookman*, *Ebony Glover*, *Ebony Baylor*, and *Ebony Utley*. I harvested the full names appearing on the first three pages of search results, using searches with other degree designations to find at least 10 full names for *Ebony*. Clicking on the link associated with *Ebony Glover* displayed an image.[8] The *Ebony Glover* in this study appeared black.

Similarly, search results for *Jill PhD* listed professionals whose first name is *Jill*. Visiting links yielded Web pages with more information about each person. For example, *Jill Schneider*'s Web page had an image showing that she is white.[14]

PeekYou searches were used to harvest a sample of full names of netizens having racially associated first names. The website peekyou.com compiles online and offline information on individuals—thereby connecting residential information with Facebook and Twitter users, bloggers, and others—then assigns its own rating to reflect the size of each person's online footprint. Search results from peekyou.com list people having the highest score first, and include an image of the person.

A PeekYou search of *Ebony* listed *Ebony Small*, *Ebony Cams*, *Ebony King*, *Ebony Springer*, and *Ebony Tan*. A PeekYou search for *Jill* listed *Jill Christopher*, *Jill Spivack*, *Jill English*, *Jill Pantozzi*, and *Jill Dobson*. After harvesting these and other full names, I reported the race of the person if discernible.

Figure 3. Image search results for first names Latanya, Latisha, Kirsten, and Jill.
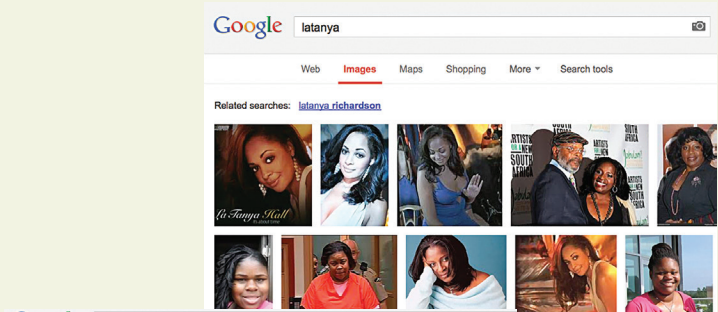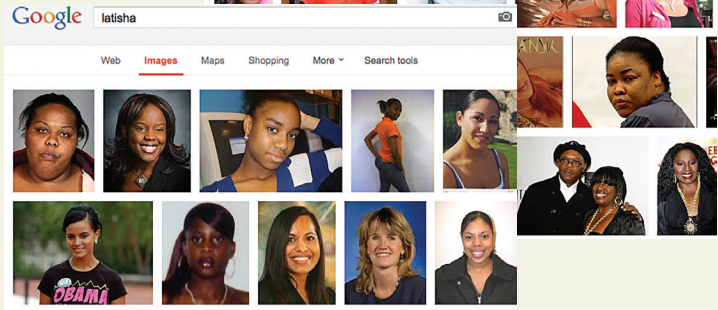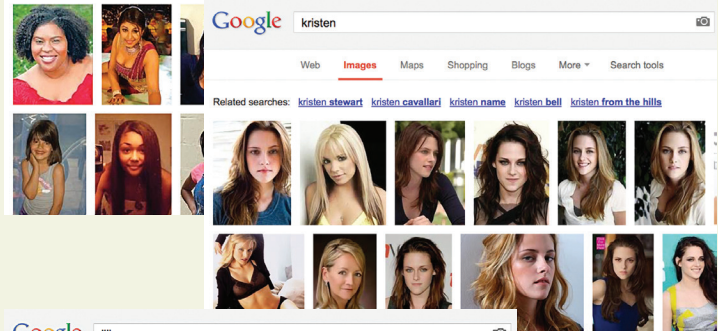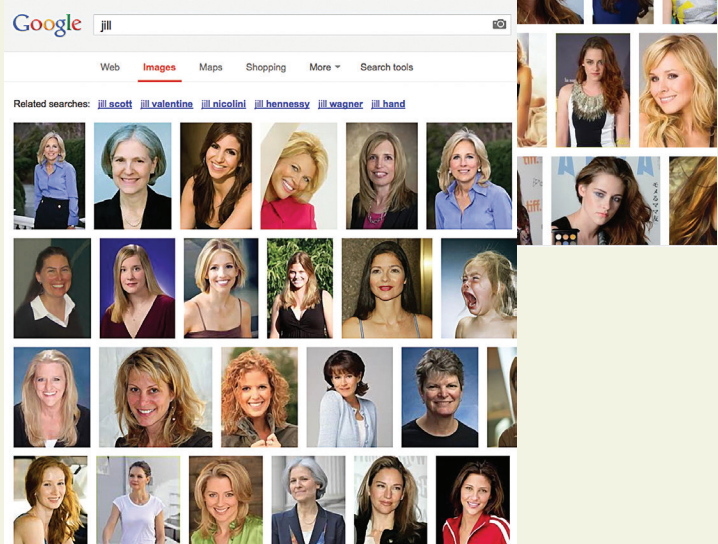
Armed with the approach just described, I harvested 2,184 racially associated full names of people with an online presence from September 24 through October 22, 2012. Most images associated with black-identifying names were of black people (88%), and an even greater percentage of images associated with white-identifying names were of white people (96%).[15]

Google searches of first names and degree designations were not as productive as first name lookups on PeekYou. On Google, white male

names, *Cody*, *Connor*, *Tanner*, and *Wyatt* retrieved results with those as last names rather than first names; the black male name, *Kenya*, was confused with the country; and black names *Aaliyah*, *Deja*, *Diamond*, *Hakim*, *Malik*, *Marquis*, *Nia*, *Precious*, and *Rasheed* retrieved fewer than 10 full names. Only *Diamond* posed a problem with PeekYou searches, seemingly confused with other online entities. *Diamond* was therefore excluded from further consideration.

Some black first names had perfect predictions (100%): *Aaliyah*, *DeAndre*, *Imani*, *Jermaine*, *Lakisha*, *Latoya*, *Malik*, *Tamika*, and *Trevon*. The worst predictors of blacks were *Jamal* (48%) and *Leroy* (50%). Among white first names, 12 of 31 names made perfect predictions: *Brad*, *Brett*, *Cody*, *Dustin*, *Greg*, *Jill*, *Katelyn*, *Katie*, *Kristen*, *Matthew*, *Tanner*, and *Wyatt*; the worst predictors of whites were *Jay* (78%) and *Brendan* (83%). These findings strongly support the use of these names as racial indicators in this study.

Sixty-two full names appeared in the list twice even though the people were not necessarily the same. No name appeared more than twice. Overall, Google and PeekYou searches tended to yield different names.

Ad Delivery
With this list of names suggestive of race, I was ready to test which ads appear when these names are searched. To do this, I examined ads delivered on two sites, Google.com and Reuters.com, in response to searches of each full name, once at each site. The browser's cache and cookies were cleared before each search, and copies of Web pages received were preserved. Figures 1, 2, 5, and 6 provide examples.

From September 24 through October 23, 2012, I searched 2,184 full names on Google.com and Reuters.com. The searches took place at different times of day, different days of the week, with different IP and machine addresses operating in different parts of the United States using different browsers. I manually searched 1,373 of the names and used automated means[17] for the remaining 812 names. Here are nine observations.

1. *Fewer ads appeared on Google.com than Reuters.com*—about five times

fewer. When ads did appear on Google.com, typically only one ad showed, compared with three ads routinely appearing on Reuters.com. This suggests Google may be sensitive to the number of ads appearing on Google.com.

2. *Of 5,337 ads captured, 78% were for government-collected information (public records) about the person whose name was searched.* Public records in the U.S. often include a person's address, phone number, and criminal history. Of the more than 2,000 names searched, 78% had at least one ad for public records about the person being searched.

3. *Four companies had more than half of all the ads captured.* These companies were Instant Checkmate, PublicRecords (which is owned by Intelius), PeopleSmart, and PeopleFinders, and all their ads were selling public records. Instant Checkmate ads appeared more than any other: 29% of all ads. Ad distribution was different on Google's site; Instant Checkmate still had the most ads (50%), but Intelius.com, while not in the top four overall, had the second most ads on Google.com. These companies dominate the advertising space for online ads selling public records.

4. *Ads for public records on a person appeared more often for those with black-associated names than white-associated names, regardless of company.* PeopleSmart ads appeared disproportionately higher for black-identifying names—41% as opposed to 29% for white names. PublicRecords ads appeared 10% more often for those with black first names than white. Instant Checkmate ads displayed only slightly more often for black-associated names (2% difference). This is an interesting finding and it spawns the question: Public records contain information on everyone, so why more ads for black-associated names?

5. *Instant Checkmate ads dominated the topmost ad position.* They occupied that spot in almost half of all searches on Reuters.com. This suggests Instant Checkmate offers Google more money or has higher quality scores than do its competitors.

6. *Instant Checkmate had the largest percentage of ads in virtually every first-name category, except for Kristen, Connor, and Tremayne.* For those names, Instant Checkmate had uncharacteristically fewer ads (less than 25%). Pub-

> **Of the more than 2,000 names searched, 78% had at least one ad for public records about the person being searched.**

licRecords had ads for 80% of names beginning with *Tremayne*, and *Connor*, and 58% for *Kristen*, compared to 20% and less for Instant Checkmate. Why the underrepresentation in these first names? During a conference call with company's representatives, they asserted that Instant Checkmate gave the same ad text to Google for groups of last names (not first names).

7. *Almost all ads for public records included the name of the person, making each ad virtually unique, but beyond personalization, the ad templates showed little variability.* The only exception was Instant Checkmate. Almost all People-Finder ads appearing on Reuters.com used the same personalized template. PublicRecords used five templates and PeopleSmart seven, but Instant Checkmate used 18 different ad templates on Reuters.com. Figure 4 enumerates ad templates for frequencies of 10 or more for all four companies (replace fullname with the person's first and last name).

While Instant Checkmate's competitors also sell criminal history information, only Instant Checkmate ads used the word arrest.

8. *A greater percentage of Instant Checkmate ads using the word "arrest" appeared for black-identifying first names than for white first names.* 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.)

The highest percentage of neutral ads (where the word *arrest* does not appear in ad text) on Reuters.com were those for *Jill* (77%) and *Emma* (75%), both white-identifying names. Names receiving the highest percentage of ads with *arrest* in the text were *Darnell*

(84%), *Jermaine* (81%), and *DeShawn* (86%), all black-identifying first names. Some names appeared counter to this pattern: *Dustin*, a white-identifying name, generated *arrest* ads in 81% of searches; and *Imani*, a black-identifying name, resulted in neutral ads in 75% of searches.

9. *Discrimination results on Google's site were similar, but, interestingly, ad text and distributions were different.* While the same neutral and *arrest* ads having dominant appearances on Reuters.com also appeared frequently on Google.com, Instant Checkmate ads on Google included an additional 10 templates, all using the word *criminal* or *arrest*.

More than 400 Instant Checkmate ads appeared on Google, and 90% of these were suggestive of *arrest*, regardless of race. Still, a greater percentage of Instant Checkmate ads suggestive of arrest displayed for black-associated first names than for whites. Of the 366

ads that appeared for black-identifying names, 92% were suggestive of arrest. Far fewer ads displayed for white-identifying names (66 total), but 80% were suggestive of arrest. This difference in the ratios 92 and 80 is statistically significant, with less than a 1% probability that the data can be explained by chance (chi-square test: $X^2(1)=7.71$, $p < 0.01$). The EEOC's adverse impact test for measuring discrimination is 40%, so if this were employment, a charge of discrimination might result. (The adverse impact test gives 100-92=8 and 100-80=20; dividing 8 by 20 equals 40.)

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.

Figure 4. Template for ads for public records on Reuters for frequencies less than 10. Full list is available.[15]

instantcheckmate		Peoplesmart	
382	**Located: *fullname*** Information found on *fullname* *fullname* found in database.	87	**We found: *fullname*** 1) Get Aisha's Background Report 2) Current Contact Info—Try Free!
96	**We found *fullname*** Search Arrests, Address, Phone, etc. Search records for *fullname*.	105	**We found: *fullname*** 1) Contact *fullname*—Free Info! 2) Current Address, Phone & More.
40	**Background of *fullname*** Search Instant Checkmate for the Records of *fullname*	348	**We found: *fullname*** 1) Contact *fullname*—Free Info! 2) Current Phone, Address & More.
17	***fullname*'s Records** 1) Enter Name and State. 2) Access Full Background Checks Instantly.		
195	***fullname*: Truth** Arrests and Much More. Everything About *fullname*		**Publicrecords**
		570	***fullname*** Public Records Found For: *fullname*. View now.
67	***fullname* Truth** Looking for *fullname*? Check *fullname*'s Arrests	128	***fullname*** Public Records Found For: *fullname*. Search now.
176	***fullname*, Arrested?** 1) Enter Name and State. 2) Access Full Background Checks Instantly.	13	**Records: *fullname*** Database of all lastname's in the Country. Search now.
55	***fullname* Located** Background Check, Arrest Records, Phone, & Address. Instant, Accurate	56	***fullname*** We have Public Records For: *fullname*. Search Now.
62	**Looking for *fullname*?** Comprehensive Background Report and More on *fullname*		**Peoplefinders**
		523	**We found *fullname*** Current Address, Phone and Age. Find *fullname*, Anywhere.

Figure 5. Senator Claire McCaskill's campaign ad appeared next to an ad using the word "arrest."

Ads by Google

Support Claire McCaskill
Claire McCaskill Fights for Middle Class America. Give to Her Campaign
www.clairemccaskill.com/

Claire McCaskill: Truth
Arrests and Much More. Everything About Claire McCaskill
www.instantcheckmate.com/

Donate to the DSCC Now
Only You Can Keep the Senate Blue In 2012. Give to the DSCC Today!
www.dscc.org/Donate
Join the DSCC! - Donate !

Figure 6. An assortment of ads appearing for Latisha Smith.

Ads by Google
Latisha Smith Located
Background Check, Arrest Records, Phone, & Address. Instant, Accurate
www.instantcheckmate.com/
(a)

Ads by Google
Latisha smith: Truth
Arrests and Much More. Everything About Latisha smith
www.instantcheckmate.com/
(b)

Ads related to Latisha Smith ⓘ
Latisha Smith, Arrested?
www.instantcheckmate.com/
1) Enter Name and State. 2) Access Full Background Checks Instantly.
(c)

Ad related to Latisha Smith ⓘ
Latisha Smith, Arrested? - 1) Enter Name and State.
www.instantcheckmate.com/
2) Access Full Background Checks Instantly.
(d)

These findings reject the hypothesis that no difference exists in the delivery of ads suggestive of an arrest record based on searches of racially associated names.

Additional Observations

The people behind the names used in this study are diverse. Political figures included Maryland State Representatives Aisha Braveboy (arrest ad) and Jay Jacobs (neutral ad); Jill Biden (neutral ad), wife of U.S. Vice President Joe Biden; and Claire McCaskill, whose campaign ad for the U.S. Sen-ate in Missouri appeared alongside an Instant Checkmate ad using the word *arrest* (Figure 5). Names mined from academic websites included graduate students, staff, and accomplished academics, such as Amy Gutmann, president of the University of Pennsylvania. Dustin Hoffman (arrest ad) was among names of celebrities used. A smorgasbord of athletes appeared, from local to national fame (assorted neutral and arrest ads). The youngest person whose name was used in the study was a missing 11-year-old black girl.

More than 1,100 of the names harvested for this study were from PeekYou, with scores estimating the name's overall presence on the Web. As expected, celebrities get the highest scores of 10s and 9s. Only four names used here had a PeekYou score of 10, and 12 had a score of 9, including Dustin Hoffman. Only two ads appeared for these high-scoring names; an abundance of ads appeared across the remaining spectrum of PeekYou scores. We might presume that the bid price needed to display an ad is greater for more popular names with higher PeekYou scores. Knowing that very few high-scoring people were in the study and that ads appeared across the full spectrum of PeekYou scores reduces concern about variations in bid prices.

Different Instant Checkmate ads sometimes appeared for the same person. About 200 names had Instant Checkmate ads on both Reuters.com and Google.com, but only 42 of these names received the same ad. The other 82% of names received different ads across the two sites. At most, three distinct ads appeared across Reuters.com and Google.com for the same name. Figure 6 shows the assortment of ads appearing for *Latisha Smith*. Having different possible ad texts for a name reminds us that while Instant Checkmate provided the ad texts, Google's technology selected among the possible texts in deciding which to display. Figure 6 shows ads both suggestive of arrest and not, though more ads appear suggestive of arrest than not.

More About the Problem

Why is this discrimination occurring? Is Instant Checkmate, Google, or society to blame? We do not yet know. Google understands that an advertiser may not know which ad copy will work best, so the advertiser may provide multiple templates for the same search string, and the "Google algorithm" learns over time which ad text gets the most clicks from viewers. It does this by assigning weights (or probabilities) based on the click history of each ad. At first, all possible ad texts are weighted the same and are equally likely to produce a click. Over time, as people tend to click one ad copy over others, the weights change, so the ad text getting the most clicks eventually displays more frequently.

Did Instant Checkmate provide ad templates suggestive of arrest disproportionately to black-identifying names? Or did Instant Checkmate provide roughly the same templates evenly across racially associated names but users clicked ads suggestive of arrest more often for black-identifying names? As mentioned earlier, during a conference call with the founders of Instant Checkmate and their lawyer, the company's representatives asserted that Instant Checkmate gave the same ad text to Google for groups of last names (not first names) in its database; they expressed no other criteria for name and ad selection.

This study is a start, but more research is needed. To preserve research opportunities, I captured additional results for 50 hits on 2,184 names across 30 Web sites serving Google Ads to learn the underlying distributions of ad occurrences per name. While analyzing the data may prove illuminating, in the end the basic message presented in this study does not change: there is discrimination in delivery of these ads.

Technical Solutions

How can technology solve this problem? One answer is to change the quality scores of ads to discount for unwanted bias. The idea is to measure real-time bias in an ad's delivery and then adjust the weight of the ad accordingly at auction. The general term for Google's technology is *ad exchange*. This approach generalizes to other ad exchanges (not just Google's); integrates seamlessly into the way ad exchanges operate, allowing minimal modifications to harmonize ad deliveries with societal norms; and, works regardless of the cause of the discrimi-

nation—advertiser bias in placing ads or society bias in selecting ads.

Discrimination, however, is at the heart of online advertising. Differential delivery is the very idea behind it. For example, if young women with children tend to purchase baby products and retired men with bass boats tend to purchase fishing supplies, and you know the viewer is one of these two types, then it is more efficient to offer ads for baby products to the young mother and fishing rods to the fisherman, not the other way around.

On the other hand, not all discrimination is desirable. Societies have identified groups of people to protect from specific forms of discrimination. Delivering ads suggestive of arrest much more often for searches of black-identifying names than for white-identifying names is an example of unwanted discrimination, according to American social and legal norms. This is especially true because the ads appear regardless of whether actual arrest records exist for the names in the company's database.

The good news is that we can use the mechanics and legal criteria described earlier to build technology that distinguishes between desirable and undesirable discrimination in ad delivery. Here I detail the four key components:

1. *Identifying Affected Groups.* A set of predicates can be defined to identify members of protected and comparison groups. Given an ad's search string and text, a predicate returns *true* if the ad can impact the group that is the subject of the predicate and returns *false* otherwise. Statistics of baby names can identify first names for constructing race and gender groups and last names for grouping some ethnicities. Special word lists or functions that report degree of membership may be helpful for other comparisons.

In this study, ads appeared on searches of full names for real people, and first names assigned to more black or white babies formed groups for testing. These *black* and *white* predicates evaluate to *true* or *false* based on the first name of the search string.

2. *Specifying the Scope of Ads to Assess.* The focus should be on those ads capable of impacting a protected group in a form of discrimination prohibited by law or social norm. Protec-

Discrimination is at the heart of online advertising. Differential delivery is the very idea behind it.

tion typically concerns the ability to give or withhold benefits, facilities, services, employment, or opportunities. Instead of lumping all ads together, it is better to use search strings, ad texts, products, or URLs that display with ads to decide which ads to assess.

This study assessed search strings of first and last names of real people, ads for public records, and ads having a specific display URL (instantcheckmate.com), the latter being the most informative because the adverse ads all had the same display URL.

Of course, the audience for the ads is not necessarily the people who are the subject of the ads. In this study, the audience is a person inquiring about the person whose name is the subject of the ad. This distinction is important when thinking about the identity of groups that might be impacted by an ad. Group membership is based on the ad's search string and text. The audience may resonate more with a distinctly positive or negative characterization of the group.

3. *Determining Ad Sentiment.* Originally associated with summarizing product and movie reviews, sentiment analysis is an area of computer science that uses natural-language processing and text analytics to determine the overall attitude of a writing.[13] Sentiment analysis can measure whether an ad's search string and accompanying text has positive, negative, or neutral sentiment. A literature search does not find any prior application to online ads, but a lot of research has been done assessing sentiment in social media (sentiment140.com, for example, reports the sentiment of tweets, which like advertisements have limited words).

In this study, ads containing the word *arrest* or *criminal* were classified as having negative sentiment; ads without those words were classified as neutral.

4. *Testing for Adverse Impact.* Consider a table where columns are comparative groups, rows are sentiment, and values are the number of ad impressions (the number of times an ad appears, though the ad is not necessarily clicked). Ignore neutral ads. Comparing the percentage of ads having the same positive or negative sentiment across groups reveals the degree to which one group may be impacted more or less by the ad's sentiment.

Table 2. Negative and neutral sentiments of black and white groups.

	Black		White	
Negative	291	60%	308	48%
Neutral	197	40%	330	52%
Positive				
Totals	**488**		**638**	

A chi-square test can determine statistical significance, and the adverse impact test used by the EEOC and the U.S. Department of Labor can alert whether in some circumstances legal risks may result.

In this study the groups are black and white, and the sentiments are negative and neutral. Table 2 shows a summary chart. Of the 488 ads that appeared for the black group, 291 (or 60%) had negative sentiment. Of the 638 ads displayed for the white group, 308 (or 48%) had negative sentiment. The difference is statistically significant ($X^2(1)=14.32$, $p < 0.001$) and has an adverse impact measure of (40/52), or 77%.

An easy way of incorporating this analysis into an ad exchange is to decide which bias test is critical (for example, statistical significance or adverse impact test) and then factor the test result into the quality score for the ad at auction. For example, if we were to modify the ad exchange not to display any ad having an adverse impact score of less than 80, which is the EEOC standard, then arrest ads for blacks would sometimes appear, but would not be overly disproportionate to whites, regardless of advertiser or click bias.

Though this study served as an example throughout, the approach generalizes to many other forms of discrimination and combats other ways ad exchanges may foster discrimination.

Suppose female names tend to get neutral ads such as "Buy now," while male names tend to get positive ads such as "Buy now. 50% off!" Or suppose black names tend to get neutral ads such as "Looking for Ebony Jones," while white names tend to get positive ads such as "Meredith Jones. Fantastic!" Then the same analysis would suppress some occurrences of the positive ads so as not to foster a discriminatory effect.

This approach does not stop the appearance of negative ads for a store placed by a disgruntled customer or ads placed by competitors on brand names of the competition, unless these are deemed to be protected groups.

Nonprotected marketing discrimination can continue even to protected groups. For example, suppose search terms associated with blacks tend to get neutral ads for some music artists, while those associated with whites tend to get neutral ads for other music artists. All ads would appear regardless of the disproportionate distribution because the ads are not subject to suppression.

As a final example, this approach allows everyone to be negatively impacted as long as the impact is approximately the same. Suppose all ads for public records on all names, regardless of race, were equally suggestive of arrest and had almost the same number of impressions; then no ads suggestive of arrest would be suppressed.

Computer scientist Cynthia Dwork and her colleagues have been working on algorithms that assure racial fairness.[4] Their general notion is to ensure similar groups receive similar ads in proportions consistent with the population. Utility is the critical concern with this direction because not all forms of discrimination are bad, and unusual and outlier ads could be unnecessarily suppressed. Still, their research direction looks promising.

In conclusion, this study demonstrates that technology can foster discriminatory outcomes, but it also shows that technology can thwart unwanted discrimination.

Acknowledgments

References
1. Barker R. *The Social Work Dictionary* (5th ed.). NASW Press, Washington, DC,ss, 2003.
2. Bertrand, M. and Mullainathan, S. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. NBER Working Paper No. 9873, 2003; http://www.nber.org/papers/w9873.
3. *Central Hudson Gas & Electric Corp. v. Public Service Commission of New York.* Supreme Court of the United States, 447 U.S. 557, 1980.
4. Dwork, C., Hardt, M., et al. 2011. Fairness through awareness. arXiv:1104.3913; http://arxiv.org/abs/1104.3913.
5. Equal Employment Opportunity Commission. Consideration of arrest and conviction records in employment decisions under Title VII of the Civil Rights Act of 1964. Washington, DC, 915.002, 2012. http://www.eeoc.gov/laws/guidance/arrest_conviction.cfm.
6. Equal Employment Opportunity Commission. Uniform guidelines on employee selection procedures. Washington, DC, 1978.
7. Fryer, R. and Levitt, S. The causes and consequences of distinctively black names. *The Quarterly Journal of Economics* 59, 3 (2004); http://pricetheory.uchicago.edu/levitt/Papers/FryerLevitt2004.pdf.
8. Glover, E.; http://www.physiology.emory.edu/FIRST/ebony2.htm (archived at http://foreverdata.org/onlineads).
9. Google AdSense; http://google.com/adsense.
10. Google. Google announces first quarter 2011 financial results; http://investor.google.com/earnings/2011/Q1_google_earnings.html.
11. Harris, P. and Keller, K. Ex-offenders need not apply: The criminal background check in hiring decisions. *Journal of Contemporary Criminal Justice* 21, 1 (2005), 6-30.
12. Panel on Methods for Assessing Discrimination, National Research Council. Measuring racial discrimination. National Academy Press, Washington, DC, 2004.
13. Pang, B. and Lee, L. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics* (2004).
14. Schneider, J. http://www.lehigh.edu/bio/jill.html (Archived at http://foreverdata.org/onlineads).
15. Sweeney, L. Discrimination in online ad delivery (2013). (For details, see full technical report at http://ssrn.com/abstract=2208240. Data, including Web pages and ads, archived at http://foreverdata.org/onlineads).
16. U.S. Commission on Civil Rights. Racism in America and how to combat it. Washington, DC, 1970.
17. WebShot Command Line Server Edition. Version 1.9.1.1; http://www.websitescreenshots.com/.

Latanya Sweeney (latanya@fas.harvard.edu) is professor of government and technology in residence at Harvard University. She creates and uses technology to assess and solve societal, political, and governance problems and teaches others how to do the same. She is also founder and director of the Data Privacy Lab at Harvard.