acmqueue

Discrimination in Online Ad Delivery

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

Latanya Sweeney

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 many more times would an ad have to appear adversely affecting one racial group for it to be considered discrimination? Is online activity so ubiquitous that computer scientists have to think about societal consequences such as structural racism in technology design? If so, how is this technology to be built? 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, a scholarship, an appointment, a promotion, or a new job, or maybe you are in a position of trust, such as a professor, a physician, a banker, a judge, a manager, or a volunteer. Perhaps you are completing a rental application, selling goods, applying for a loan, joining a social club, making new friends, dating, or engaged in any one of hundreds of circumstances for which someone wants to learn more about you online. Appearing alongside your list of 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.

Job applications frequently include questions such as: Have you ever been arrested? Have you ever been charged with a crime? Other than a traffic ticket, have you been convicted of a crime? Employers ask these questions to establish trustworthiness. Because others often equate a criminal record with not being reliable or honest, protections exist for those having criminal records.

If an employer disqualifies a job applicant based solely upon information indicating an arrest record, the company may face legal consequences. The U.S. EEOC (Equal Employment Opportunity Commission) 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. Guidance issued in 1973 extended protections to people with criminal records.^{5,11} Title VII does not prohibit employers from obtaining criminal background information. Certain uses of criminal information, however, such as a blanket policy or practice of excluding applicants or disqualifying employees 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. For example, suppose a company laid off comparable black and white workers at the same rate—25 percent of blacks and 25 percent of whites—then the ratio, 25 divided by 25, would be 100 percent. If the ratio is less than 80 percent, 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, even when no one with your name has been arrested? Title VII does not apply unless you have an arrest record and can prove the potential employer routinely uses ads or information from the company sponsoring the ads, and the result has an inappropriate chilling effect on hiring applicants with criminal records.

The advertiser may argue the ads are commercial free speech—a constitutional right to display the ad associated with your name. The First Amendment of the U.S. Constitution protects advertising. In a landmark decision, the U.S. Supreme Court set out a test for assessing government 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?

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 the 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 . . . Racism is not just a matter of attitudes; actions and institutional structures can also be a form of racism.”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, because discrimination depends on the ability to give or withhold benefits, facilities, services, opportunities, etc., 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

Racism can result, even if not intentional, and online activity now may be so ubiquitous that computer scientists have to think about societal consequences such as structural racism in the technology they design. These considerations frame the big picture, the relevant legal, societal, and technical landscape in which this exploration resides. Now we turn to the exploration itself: whether online ads suggestive of arrest records appear more often for one racial group than another among a sample of racially associated names. Then, we examine the role technology might play in combating this problem if evidence of the pattern exists.

THE PATTERN

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

This study begins with the assumption that personalized ads suggestive of arrest records do not differ by race. We did this by carefully constructing the scientifically best instance of the pattern—one with names shown to be racially identifying and pseudo-randomly selected.

Earlier this year, a Google search for *Latanya Farrell*, *Latanya Sweeney*, and *Latanya Lockett* yielded the ads and criminal reports shown in figure 1. The ads appeared on Google.com (figure 1a,1c,1e) and on a news Web site, Reuters.com, to which Google supplies ads (figure 1c, bottom). All the ads in question linked to instantcheckmate.com (figure 1b,1d,1f). The first ad implied Latanya Farrell may have been arrested. Was she? Clicking on the link and paying the requisite subscription fee revealed

FIGURE 1

Sample Ads and Criminal Reports

A

Ads related to latanya farrell

[Latanya Farrell, Arrested?](#)www.instantcheckmate.com/

1) Enter Name and State. 2) Access Full Background Checks Instantly.

[Latanya Farrell](#)www.publicrecords.com/Public Records Found For: **Latanya Farrell**. View Now.

C

Ad related to latanya sweeney

[Latanya Sweeney Truth](#)www.instantcheckmate.com/Looking for **Latanya Sweeney**? Check **Latanya Sweeney's** Arrests.

D

F

	Name	County and State	Offense	View Details
1	Latanya Lockett	CA Contra Costa County (San)	Criminology	View Details
2	Latanya Lockett	CA Contra Costa County	Criminology	View Details
3	Latanya Lockett	CA Orange Superior Court	Criminology	View Details
4	Latanya Lockett	CA Orange Superior Court	Criminology	View Details

B

C

Ads by Google

[Latanya Sweeney, Arrested?](#)

1) Enter Name and State. 2) Access Full Background Checks Instantly.

www.instantcheckmate.com/[Latanya Sweeney](#)Public Records Found For: **Latanya Sweeney**. View Now.www.publicrecords.com/[La Tanya](#)

Search for La Tanya Look Up Fast Results now!

www.ask.com/La+Tanya

E

Ads related to latanya lockett

[We Found:Tanya Lockett](#)www.peoplesmart.com/1) Get **Tanya Lockett's** Info - Try Free! 2) Current Phone, Address & More.[Latanya Lockett, Arrested?](#)www.instantcheckmate.com/

1) Enter Name and State. 2) Access Full Background Checks Instantly.

[Latanya Lockett Found](#)www.whitepages.com/Latanya+Lockett

Don't Pay for Info that's Free, Get Address, Phone, Photos, & More!

Name Popularity & Facts - Neighbor Search - Reverse Phone Lookup

that the company had no arrest record for her (figure 1b). There is no arrest record for *Latanya Sweeney* either, but there is for *Latanya Lockett*.

In comparison, searches for *Kristen Haring*, *Kristen Sparrow*, and *Kristen Lindquist* did not yield any instantcheckmate.com ads (figure 2a, 2c, and 2e), even though the company's database reported having records for all three names and arrest records for *Kristen Sparrow* and *Kristen Lindquist* (figure 2d and 2f).

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 3 shows the ads and criminal reports for these three names appearing on Google.com (figure 1c, 1e) and Reuters.com (figure 1a). Criminal reports came from instantcheckmate.com (figure 1b, 1d, 1f).

FIGURE 2

Sample Ads and Criminal Reports that Include the First Name,"Kristen"

A

Ads by Google

We Found: Kristen Haring

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

www.peoplesmart.com/Kristen

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

Kristen Haring

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

B

KRISTEN HARING

1108 Island Ave
San Diego, CA 92108
DOB: Oct 21, 1987 (35 years old)

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

Location
Current address history and reported data, maps, etc.

Related Persons
Known family members, business associates, coworkers, etc.

Marriage / Divorce
Marriage and divorce records on file.

Criminal History
Arrest records, speeding tickets, mugshots, etc.

Licenses
Auto licenses, DEA licenses, Other Licenses, etc.

Sex Offenders
Sex offenders living near Kristen Haring's primary location.

Criminal History

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 clear state here should not be interpreted as a guarantee that Kristen Haring 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.			

C

Ads by Google

We Found: Kristen Sparrow

1) Contact **Kristen Sparrow** - Free Info! 2) Current Phone, Address & More.

www.peoplesmart.com/

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

Kristen Sparrow

Public Records Found For: **Kristen Sparrow**. View Now.
www.publicrecords.com/

D

KRISTEN SPARROW

2411 Greenwood St
San Francisco, CA 94123
DOB: Nov 30, 1993 (29 years old)

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

Location
Current address history and reported data, maps, etc.

Related Persons
Known family members, business associates, coworkers, etc.

Marriage / Divorce
Marriage and divorce records on file.

Criminal History
Arrest records, speeding tickets, mugshots, etc.

Licenses
Auto licenses, DEA licenses, Other Licenses, etc.

Sex Offenders
Sex offenders living near Kristen Sparrow's primary location.

Criminal History

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 clear state here should not be interpreted as a guarantee that Kristen Sparrow 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
1. Kristen Tracy Sparrow	CA San Mateo County Superior Court	Criminal Traffic	View Details

E

Ads by Google

Kirsten Lindquist

Get **Kirsten Lindquist** Find **Kirsten Lindquist**
www.ask.com/Kirsten+Lindquist

We Found: Kristen Lindquist

1) Contact **Kristen Lindquist** - Free Info! 2) Current Phone, Address & More.

www.peoplesmart.com/

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

Kristen Lindquist

Public Records Found For: **Kristen Lindquist**. View Now.
www.publicrecords.com/

F

KRISTEN LINDQUIST

122 Hawthorne Ct
Charlotte, NC 28204
DOB: Nov 30, 1984 (39 years old)

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

Location
Current address history and reported data, maps, etc.

Related Persons
Known family members, business associates, coworkers, etc.

Marriage / Divorce
Marriage and divorce records on file.

Criminal History
Arrest records, speeding tickets, mugshots, etc.

Licenses
Auto licenses, DEA licenses, Other Licenses, etc.

Sex Offenders
Sex offenders living near Kristen Lindquist's primary location.

Criminal History

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 clear state here should not be interpreted as a guarantee that Kristen Lindquist 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
1. Kristen Marie Lindquist	Individual NC courts	Criminal Traffic	View Details
2. Kristen Marie Lindquist	NC Admin Office of Courts demographic criminal	Criminal Traffic	View Details
3. Kristen Marie Lindquist	NC Admin Office of Courts demographic criminal	Criminal Traffic	View Details

Finally, we considered a proxy for race associated with these names. Figure 4 shows a racial distinction in the Google images that appear for image searches of *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*.

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

FIGURE 3

Sample Ads and Criminal Reports for Names that Include the First Name “Jill”

A

Ads by Google

Located: Jill Foley
Information found on **Jill Foley** **Jill Foley** found in database.
www.instantcheckmate.com/

Macy's @ Wedding Registry
Official Site. Create a Registry or Buy a Wedding Gift at Macy's!
www.macys.com/Registry/Wedding
macys.com is rated ★★★★★ (82 reviews)

Dr. Jill Foley
View Credentials, Malpractice, Bio, Ratings, Reviews & Background Now!
www.lifescript.com/MD

C

Ads related to Jill Schneider

Jill Schneider Art
www.posters2prints.com/
Custom Frame Prints and Canvas. Shop Now, SAVE Big + Free Shipping!

We Found Jill Schneider
www.intelius.com/
Current Phone, Address, Age & More. Instant & Accurate **Jill Schneider**
10,266 people +1'd this page
Reverse Lookup - Reverse Cell Phone Directory - Date Check - Property Records

Located: Jill Schneider
www.instantcheckmate.com/
Information found on **Jill Schneider** **Jill Schneider** found in database.

E

Ad related to Jill James

Located: Jill James
www.instantcheckmate.com/
Information found on **Jill James** **Jill James** found in database.

B

JILL FOLEY
350 Prospect Rd.
North Palm Beach, FL 33408
DOB: Jul 17, 1941 (64 years old)

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

Location
Current address history and related facts, maps, etc.

Related Persons
Recent family members, business associates, acquaintances, etc.

Marriage / Divorce
Marriage and divorce records on file.

Criminal History
Arrest records, spending habits, mugshots, etc.

Licenses
Real estate, DEA licenses, other licenses, etc.

Sex Offenders
Sex offenders living near Jill Foley's primary location.

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 Jill Foley 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
1. Jill Lee Foley	FL, Pinellas County	Concealment	View Details
2. Jill Lee Foley	FL, Pinellas County	Concealment	View Details

D

JILL SCHNEIDER
1501 75th St.
Pewaukee, WI 53191
DOB: Mar 31, 1989 (35 years old)

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

Location
Current address history and related facts, maps, etc.

Related Persons
Recent family members, business associates, acquaintances, etc.

Marriage / Divorce
Marriage and divorce records on file.

Criminal History
Arrest records, spending habits, mugshots, etc.

Licenses
Real estate, DEA licenses, other licenses, etc.

Sex Offenders
Sex offenders living near Jill Schneider's primary location.

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 Jill Schneider 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
1. Jill E. Schneider	WI Admin Office of Courts(CD)	Concealment	View Details
2. Jill E. Schneider	WI Admin Office of Courts(CD)	Concealment	View Details
3. Jill E. Schneider	WI Admin Office of Courts(CD)	Concealment	View Details
4. Jill E. Schneider	WI Admin Office of Courts(CD)	Concealment	View Details

F

JILL JAMES
600 Seward St.
Cary, NC 27513
DOB: May 31, 1952 (54 years old)

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

Location
Current address history and related facts, maps, etc.

Related Persons
Recent family members, business associates, acquaintances, etc.

Marriage / Divorce
Marriage and divorce records on file.

Criminal History
Arrest records, spending habits, mugshots, etc.

Licenses
Real estate, DEA licenses, other licenses, etc.

Sex Offenders
Sex offenders living near Jill James's primary location.

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 Jill James 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
1. Jill B. James	NC Admin Office of Courts demographic criminal	Concealment	View Details
2. Jill James	NC Admin Office of Courts demographic criminal	Concealment	View Details
3. Jill James	Individual NC courts	Concealment	View Details
4. Jill B. James	Individual NC courts	Concealment	View Details
5. Jill B. James	Individual NC courts	Concealment	View Details
6. Jill B. James	NC Admin Office of Courts demographic criminal	Concealment	View Details
7. Jill Kaly James	NC Admin Office of Courts demographic criminal	Concealment	View Details
8. Jill Kaly James	Individual NC courts	Concealment	View Details
9. Jill Rosemond James	NC Admin Office of Courts demographic infraction	Concealment	View Details
10. Jill Rosemond James	NC Admin Office of Courts demographic criminal	Concealment	View Details

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 Web site such as Reuters, and Instant Checkmate in the previous examples? An overview of Google AdSense, the program that delivered the ads, explains the links between these companies.

In printed newspapers and magazines, ad space and ad content are fixed. Traditionally, everyone who reads the publication sees the same ad in the same space. Web sites are different. Online ad space, not bound by the same physical limitations, can be dynamic, with ads tailored to the reader's

FIGURE 4

Sample Face Images on Google.com

A

B

C

E

Ad related to Jill James ⓘ
Located: Jill James
www.instantcheckmate.com/
 Information found on Jill James Jill James found in database.

D

checkmate DASHBOARD EDIT ACCOUNT INFO LOGOUT

JILL SCHNEIDER
 1707 700 St
 New York, NY 10119
 DOB: Mar 31, 1989 (24 years old)

Personal
 Name, email, contacts, phone numbers, etc.

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

Related Persons
 Current family members, business associates, roommates, etc.

Marriage / Divorce
 Marriage and divorce records on file.

Criminal History
 Arrest records, pending tickets, mugshots, etc.

Licenses
 Driver licenses, DEA licenses, Other Licenses, etc.

Sex Offenders
 Sex offenders living near Jill Schneider's primary location.

Criminal History
 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 Jill Schneider 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
1. Jill E Schneider	WI Admin Office of Court(20) inspection	Conductafrica	View Details
2. Jill E Schneider	WI Admin Office of Court(20)	Conductafrica	View Details
3. Jill E Schneider	WI Admin Office of Court(20) inspection	Conductafrica	View Details
4. Jill E Schneider	WI Admin Office of Court(20)	Conductafrica	View Details

F

checkmate DASHBOARD EDIT ACCOUNT INFO LOGOUT

KRISTEN LINDQUIST
 714 Hawthorne Ln
 Charlotte, NC 28217
 DOB: Nov 30, 1984 (28 years old)

Personal
 Name, email, contacts, phone numbers, etc.

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

Related Persons
 Current family members, business associates, roommates, etc.

Marriage / Divorce
 Marriage and divorce records on file.

Criminal History
 Arrest records, pending tickets, mugshots, etc.

Licenses
 Driver licenses, DEA licenses, Other Licenses, etc.

Sex Offenders
 Sex offenders living near Kristen Lindquist's primary location.

Criminal History
 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 Kristen Lindquist 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
1. Kristen Marie Lindquist	Individual NC courts	Conductafrica	View Details
2. Kristen Marie Lindquist	NC Admin Office of Courts demographics criminal	Conductafrica	View Details
3. Kristen Marie Lindquist	NC Admin Office of Courts demographics criminal	Conductafrica	View Details

search criteria, interests, geographical location, and so on. Any two readers (or even the same reader returning to the same Web site) might view different ads.

Google AdSense is the largest provider of dynamic online advertisements, placing ads for millions of sponsors on millions of Web sites.⁹ In the first quarter of 2011, Google earned US\$2.43 billion (\$9.71 billion annualized), or 28 percent of its total revenue, 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.

When a reader enters search criteria on an enrolled Web site, Google AdSense embeds into

the page of results ads that are believed to be relevant to the search. Figures 1, 2, and 3 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 once a match occurs, and a bid of how much the sponsor is willing to pay if a reader clicks the delivered ad. (This article conflates two interacting Google programs: Google AdWords allows advertisers to specify search criteria, ad text, and bids; and Google AdSense delivers the ads to host sites.) Google operates a realtime auction across bids for the same search criteria, computing an overall “quality score” to use as the basis for the auction. The quality score includes many factors such as the past performance of the ad and characteristics of the company’s Web site.¹⁰ The ad with 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 (e.g., a maximum account total for the sponsor). The Instant Checkmate ads in figures 1, 2, and 3 often appeared first among ads, implying Instant Checkmate had the highest quality score.

A Web-site owner that wants to “host” online ads enrolls in AdSense and modifies the Web site to include special software that sends information about the current reader (e.g., search criteria) to Google; in exchange, the Web site receives corresponding ads from Google. The displayed ads have the banner “Ads by Google” when they appear on sites other than Google.com. For example, Reuters.com is an AdSense host, and entering *Latanya Sweeney* in the search bar generated a new Web page with ads delivered by Google, bearing the banner “Ads by Google” (figure 1c).

There is no cost associated with displaying an ad, but if the user actually clicks the ad, the sponsor pays the bid 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 bid to Google, and Google would split the payment with Reuters.

SEARCH CRITERIA

What search criteria did Instant Checkmate specify? Are ads randomly delivered? Do ads rely only on the first name? Will ads be delivered for made-up names? Google AdSense provides answers to these questions. Ads displayed on Google.com allow users to learn why a specific ad appeared. Clicking the circled “i” in the ad banner (e.g., figure 1c) leads to a Web page explaining the ads. Doing so for ads in figures 1 and 3 reveals that the ads appeared because the search criteria associated with the bid matched the exact first- and last-name combination searched. Because a company presumably bids on records it sells, the names would likely be the first and last names of real people.

This means that the search criteria associated with these ads have to consist of both first and last names, and the names should belong to real people.

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. Instant Checkmate is not presumed to have used such a list in placing bids, nor Google in delivering ads. Rather, the list provides a qualified sample of racially associated 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 NBER (National Bureau of Economic

Research) did a field experiment in which they provided resumes to job ads that were virtually identical, except that some of the resumes had black-identifying names and others had white-identifying names.² Their job discrimination study showed significant discrimination against black names: white names received 50 percent more callbacks for interviews, even though the resumes otherwise had identical qualifications.

The study used a correlation of 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* (William Morrow, 2006), Steven Levitt and Stephen Dubner report the top 20 whitest- and blackest-identifying girls' and boys' 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, which they correlate with the Black Power Movement.⁷ They postulate that the movement influenced how blacks perceived their identities, and they give as evidence that before the movement, names given to black and white children were not distinctly different, but after the movement distinctly black names emerged.

Similar to the job discrimination study, the list used by Fryer and Levitt 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 methods of ad delivery, we 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 job discrimination 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 4. Removing duplicates leaves a total of 63 distinct first names.

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

White Female	Black Female	White Male	Black Male
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
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
	Latanya		
	Latisha		

FULL NAMES OF REAL PEOPLE

Having a list of racially associated first names is a start, but testing ad delivery requires a real person's first and last name (full name). Web searches provide a means of locating and harvesting full names by: (1) sampling names of professionals appearing on the Web; and (2) sampling names of people active on social media sites and blogs (netizens).

Professionals often have their own Web sites or have biographical information appearing on institutional Web sites, listing titles and positions and describing prior accomplishments and current activities. Several professions, such as research, medicine, law, and business, often have degree designations (e.g., PhD, MD, JD, 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. These kinds of searches can harvest a sample of full names of professionals with racially associated first names.

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. Each Web page visited should be archived to preserve images and content.

Here are two examples from my ad-delivery 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 professional endings such as *JD*, *MD*, or *MBA* as needed to find at least 10 full names for *Ebony*. Clicking on the link associated with *Ebony Glover* provided more information about her, including 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

Harvesting names of netizens is similar but simpler than harvesting names of professionals. PeekYou searches were used to harvest a sample of full names of netizens who have racially associated first names. The Web site peekyou.com compiles online and offline information on individuals—thereby connecting residential information with Facebook and Twitter users, bloggers, and others—and assigns its own rating for the size of each person's online footprint. Search results from peekyou.com list people with the highest score first, and include an image of the person. Celebrities and public figures tend to be listed first, having the highest PeekYou scores, followed by bloggers, tweeters, and the rest.

A PeekYou search for *Ebony* found *Ebony Small*, *Ebony Cams*, *Ebony King*, *Ebony Springer*, and *Ebony Tan*. A PeekYou search for *Jill* found *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.

Using 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. Using the images associated with those names, I was able to confirm that the racially associated first names were predictive of race.¹⁵ Most images associated with black-identifying names were of black people (88 percent), and an even greater percentage of images associated with white-identifying names were of white people (96 percent).

Black names and white names were examined separately as predictors of race. The results showed that 490 images of blacks had black-associated first names, and 68 did not; 18 images of blacks had

white first names; 852 had neither black first names nor images of blacks. Similarly, 831 images of whites had white first names, 50 images of whites did not have white first names; 39 had white first names but nonwhite images, and 508 had neither white first names nor images of whites.

Google searches of first names and degree designations were not as productive as first name lookups on PeekYou. On Google, the 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 the 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 percent): *Aaliyah*, *DeAndre*, *Imani*, *Jermaine*, *Lakisha*, *Latoya*, *Malik*, *Tamika*, and *Trevon*. The worst predictors of blacks were *Jamal* (48 percent) and *Leroy* (50 percent). 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 percent) and *Brendan* (83 percent). 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, 3, 6, and 7 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, on 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 means¹⁷ for the remaining 812 names. Here are 10 observations.

1. The ads were respectfully displayed, without clutter. We have all seen Web pages where ads get in the way, dominating the page or being so closely woven into the page that you cannot distinguish the ads from the content. That's not the case here. No more than three ads ever appeared for a search on either Google.com or Reuters.com. No company's ad was listed more than once on a page, and the ads appeared in a single set within a rectangular area in the margins. Google and Reuters are respected sources of information, and displayed in this manner, the ads did nothing to take away from the Web sites; conversely, the sites and respectful placement of ads may even exalt the ads.

2. Far fewer ads appeared on Google.com than Reuters.com—about five times fewer, even when examining up to three pages of search results on Google.com. 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.

3. Of 5,337 ads captured, 78 percent were for government-collected information (public records)

about the person whose name was searched. Public records in the United States often include a person's address, phone number, criminal history, and professional and business licenses, though specifics vary among states. Of the more than 2,000 names searched, 78 percent had at least one ad for public records about the person being searched. Ads to buy a person's public record appeared for almost any name searched, but they came up on Reuters.com much more often than on Google.com.

4. 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 percent of all ads. Ad distribution was different on Google's site; Instant Checkmate still had the most ads (50 percent), 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.

5. Instant Checkmate ads dominated the topmost ad position. They occupied that spot in almost half of all searches on Reuters.com. The next closest, PublicRecords.com, rarely had the topmost spot, but most frequently appeared in the second and third positions. Appearing as the first ad so often suggests that, in general, Instant Checkmate offers Google more money or has higher quality scores than do its competitors.

6. 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 percent as opposed to 29 percent for white names. PublicRecords ads appeared 10 percent more often for black first names than for white. Instant Checkmate ads displayed only slightly more often for black-associated names (2 percent difference). This is one of those interesting findings that spawn the question: Public records contain information on everyone, so why more ads for black-associated names?

7. 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 percent). PublicRecords had ads for 80 percent of names beginning with *Tremayne*, compared with only 23 percent for Instant Checkmate. Similarly, for *Connor*, PublicRecords had 80 percent compared with 20 percent for Instant Checkmate, and for *Kristen* it was 58 percent PublicRecords versus 16 percent Instant Checkmate. Why the underrepresentation in these first names? Did Instant Checkmate avoid these names for some reason? Do these undercounts show a glitch? 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).

8. 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. For example, almost all PeopleFinder ads appearing on Reuters.com used the same personalized template ("We found *fullname*. Current Address, Phone and Age. Find *fullname*, Anywhere," where the person's first and last name replaces *fullname*). PublicRecords used five templates and PeopleSmart seven, but Instant Checkmate used 18 different ad templates on Reuters.com. Figure 5 enumerates ad texts and frequencies for all four companies (replace *fullname* with the person's first and last name).

Only Instant Checkmate ads used the word *arrest*, which appeared in eight of its 18 ad templates found on Reuters.com. While Instant Checkmate's competitors—PeopleSmart, PublicRecords, and PeopleFinders—also sell criminal history information, none of their ads included the word *arrest* or *arrested*.

FIGURE 5

Templates for Ads for Public Records on Reuters

instantcheckmate		Peoplesmart	
382	Located: <u>fullname</u> Information found on <u>fullname</u> found in database.	7	We found: <u>fullname</u> 1) Get <u>firstname</u> 's Background Report 2) Contact info & More -Try Free!
2	Located: The Person Information found on them Person found in database.	87	We found: <u>fullname</u> 1) Get Aisha's Background Report 2) Current Contact Info - Try Free!
96	We found <u>fullname</u> Search Arrests, Address, Phone, etc. Search records for <u>fullname</u> .	105	We found: <u>fullname</u> 1) Contact <u>fullname</u> -Free Info! 2) Current Address, Phone & More.
4	We found Them Search Arrests, Address, Phone, etc. Search records for <u>fullname</u> .	348	We found: <u>fullname</u> 1) Contact <u>fullname</u> -Free Info! 2) Current Phone, Address & More.
40	Background of <u>fullname</u> Search Instant Checkmate for the Records of <u>fullname</u>	1	We found <u>firstname</u> Get <u>firstname</u> in CA's Email, Address, Phone, Public Records & More Easy!
9	Background of Anyone Search Instant Checkmate for the Records of <u>fullname</u>	1	We found <u>firstname</u> In <u>lastname</u> 1)Get <u>firstname</u> 's Info – Try Now! 2)Current Phone, Address & More.
17	<u>fullname</u>'s Records 1) Enter Name and State. 2) Access Full Background Checks Instantly.	1	Looking For <u>fullname</u>? Get <u>fullname</u> 's Phone, Email Address, Public Records & More Now!
3	Anyone's Records 1) Enter Name and State. 2) Access Full Background Checks Instantly.	Publicrecords	
195	<u>fullname</u>: Truth Arrests and Much More. Everything About <u>fullname</u>	570	<u>fullname</u> Public Records Found For: <u>fullname</u> . View now.
67	<u>fullname</u> Truth Looking for <u>fullname</u> ? Check <u>fullname</u> 's Arrests	128	<u>fullname</u> Public Records Found For: <u>fullname</u> . Search now.
176	<u>fullname</u>, Arrested? 1) Enter Name and State. 2) Access Full Background Checks Instantly.	13	Records: <u>fullname</u> Database of all <u>lastname</u> 's in the Country. Search now.
2	Uh Oh, Arrested? 1) Enter Name and State. 2) Access Full Background Checks Instantly.	2	Fullname Info View Contact Information For Free Quick & Easy Search Results!
1	Found: <u>fullname</u> We have the story on <u>fullname</u> <u>fullname</u> 's arrests, relatives, etc.	56	<u>fullname</u> We have Public Records For: <u>fullname</u> . Search Now.
3	Fullname - Found Learn the truth about <u>fullname</u> Check <u>fullname</u> 's arrests & more.	Peoplefinders	
4	Research <u>fullname</u> We have details on <u>fullname</u> . <u>fullname</u> 's full background & info.	523	We found <u>fullname</u> Current Address, Phone and Age. Find <u>fullname</u> , Anywhere.
55	<u>fullname</u> Located Background Check, Arrest Records, Phone, & Address. Instant, Accurate	8	We found <u>fullname</u> 1)Get Phone/ Address/ Age Instantly! 2) Find Anyone, Anywhere for Free.
62	Looking for <u>fullname</u>? Comprehensive Background Report and More on <u>fullname</u>	2	Find <u>fullname</u> Get current and past addresses and phone numbers. Instant results!
8	Looking for People in the US? Comprehensive Background Report and More on <u>fullname</u>	1	We Found Them for Free Current Address, Phone and Age. Find <u>fullname</u> Anywhere.

9. 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 percent used *arrest* in the ad text. Of the 638 ads displayed with white-identifying names, 48 percent used *arrest*. This difference is statistically significant, with less than a 0.1 percent probability that the data can be explained by chance (chi-square test: $\chi^2(1)=14.32$, $p < 0.001$). 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 the ad text) on Reuters.com were those for *Jill* (77 percent) and *Emma* (75 percent), both white-identifying names. Names receiving the highest percentage of ads with *arrest* in the text were *Darnell* (84 percent), *Jermaine* (81 percent), and *DeShawn* (86 percent), all black-identifying first names. Some names

appeared counter to this pattern: *Dustin*, a white-identifying name, generated *arrest* ads in 81 percent of searches; and *Imani*, a black-identifying name, resulted in neutral ads in 75 percent of searches.

10. Discrimination results on Google's site were similar, but, interestingly, ad text and distributions were different. Instant Checkmate ads appearing on Google.com often used different ad text than those on Reuters.com. While the same neutral and *arrest* ads that were dominant 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*. These new templates appeared in about 20 percent of the Instant Checkmate ads on Google.

More than 400 Instant Checkmate ads appeared on Google, and 90 percent of these were suggestive of *arrest*, regardless of race. Together, these last two findings underscore other differences between ads appearing on Google's own site and those delivered by Google AdSense to Reuters. Ad text was different. Ads with the word *criminal* and not *arrest* appeared only on Google's site, and ads using either *arrest* or *criminal* appeared much more often for both races on Google.com.

Still, on Google's own site, a greater percentage of Instant Checkmate ads suggestive of arrest displayed for black-associated first names than for white-associated names. Of the 366 ads that appeared for black-identifying names, 92 percent were suggestive of arrest. Far fewer ads displayed for white-identifying names (66 total), but 80 percent were suggestive of arrest. This difference in the ratios 92 and 80 is statistically significant, with less than a 1 percent probability that the data can be explained by chance (chi-square test: $\chi^2(1)=7.71$, $p < 0.01$). The EEOC's adverse impact test for measuring discrimination is 40 percent, so in an employment situation, a charge of discrimination might result. (The adverse impact test gives $100-92=8$ and $100-80=20$; dividing 8 by 20 gives 40 percent.)

A greater percentage of Instant Checkmate ads with the word *arrest* in ad text appeared for black-identifying first names than for white-identifying first names within professional and netizen subsets, too.

This study started with the hypothesis that no difference exists in the delivery of ads suggestive of an arrest record based on searches of racially associated names. The findings reject this. A greater percentage of ads using *arrest* in their text appeared for black-identifying first names than for white-identifying first names in searches on Reuters.com, Google.com, and in subsets of the sample. On Reuters.com, which hosts Google AdSense ads, a black-identifying name was 25 percent more likely to generate an ad suggestive of an arrest record.

THREE ADDITIONAL OBSERVATIONS

The people behind the names used in this study are diverse. Political figures included State Representatives Aisha Braveboy (arrest ad) and Jay Jacobs (neutral ad) of Maryland; Jill Biden (neutral ad), wife of U.S. Vice President Joe Biden; and Claire McCaskill, whose campaign ad for the U.S. Senate in Missouri appeared alongside an Instant Checkmate ad using the word *arrest* (figure 6). Names mined from academic Web sites included graduate students, researchers, administrators, staff, and accomplished academics, such as Amy Gutmann, president of the University of Pennsylvania and chair of the U.S. Presidential Commission for the Study of Bioethical Issues. Dustin Hoffman (arrest ad) was among the celebrity names 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.

FIGURE 6

Example Ads Displayed in Response to Search of “Claire McCaskill”

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. It seems likely that the bid price needed to get an ad placed first is greater for more well-known and 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 percent of names received different ads across the two sites. Search results on Reuters.com for the 62 duplicate names that appeared in the study showed different ads for 37 of them, the same ad for seven, and no ad for 18. At most, three distinct ads appeared across Reuters.com and Google.com for the same name. Figure 7 shows the assortment of ads appearing for *Latonya Evans* and *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. In Figure 7, each name had ads both suggestive of arrest and not, though they both had more ads suggestive of arrest than not.

MORE ABOUT THE PROBLEM

Why is this discrimination occurring? Is Instant Checkmate, Google, or society to blame? We don't yet know, but navigating the terrain requires further information about the inner workings of Google AdSense. 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 presumed equally likely to produce a click. Over time, as people click one version of an ad more often than others, the weights change, so the ad text getting the most clicks

FIGURE 7

Examples of Different Ad Copy Appearing for Different Names

A

Ads by Google

[Latonya Evans, Arrested?](#)

1) Enter Name and State. 2) Access Full Background Checks Instantly.

www.instantcheckmate.com/

A

Ads by Google

[Latonya Evans's Records](#)

1) Enter Name and State. 2) Access Full Background Checks Instantly.

www.instantcheckmate.com/

A

Ad related to Latonya Evans ⓘ

[Latonya Evans's Records](#)

www.instantcheckmate.com/

Did you know **Latonya Evans's** criminal history is searchable?

B

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.

B

Ad related to Latisha Smith ⓘ

[Latisha Smith, Arrested? - 1\) Enter Name and State.](#)

www.instantcheckmate.com/

2) Access Full Background Checks Instantly.

eventually displays more frequently. This approach aligns the financial interests of Google, as the ad deliverer, with the advertiser.

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.

Google uses cloud-caching strategies to deliver ads quickly. Might these strategies create a bias toward templates previously loaded in the cloud cache? Is there a combination effect?

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 realtime 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 integrates seamlessly into the way ad exchanges operate, allowing minimal modifications to harmonize ad deliveries with societal norms; it generalizes to other ad exchanges (not just Google's); and, finally, it works regardless of the cause of the discrimination—advertiser bias in placing ads or societal 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. Key components are: (1) identifying affected groups; (2) specifying the scope of ads to assess; (3) determining ad sentiment; and (4) testing for adverse impact.

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. Protection 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 or 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 subjects 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 text.¹³ Sentiment analysis can measure whether an ad's search string and accompanying text have 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, and 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, whether or not it is 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. 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 indicate whether in some circumstances the impact may lead to legal risks.

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 percent) had negative sentiment. Of the 638 ads displayed for the white group, 308 (or 48 percent) had negative sentiment. The difference is statistically significant ($\chi^2(1)=14.32$, $p < 0.001$) and has an adverse impact measure of 40/52, or 77 percent.

An easy way of incorporating this analysis into an ad exchange is to decide which bias test is critical (e.g., statistical significance or the 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 with 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 such ads for 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 of fostering 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.

TABLE 2. Assessment of negative and neutral ads for black and white groups

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

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 would not be subject to suppression.

As a final example, this approach allows everyone to be negatively impacted as long as the impact is roughly 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 already been working on algorithms that assure racial fairness.⁴ Their general notion is to make sure 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

The author thanks Ben Edelman, Claudine Gay, Gary King, Annie Lewis, and weekly Topics in Privacy participants (David Abrams, Micah Altman, Merce Crosas, Bob Gelman, Harry Lewis, Joe Pato, and Salil Vadhan) for discussions; Adam Tanner for first suspecting a pattern; Diane Lopez and Matthew Fox in Harvard's Office of the General Counsel for making publication possible in the face of legal threats; and Sean Hooley for editorial suggestions. Data from this study is available at foreverdata.org and the IQSS Dataverse Network. Supported in part by NSF grant CNS-1237235 and a gift from Google, Inc.

REFERENCES

1. Barker R. 2003. *The Social Work Dictionary* (5th ed.). Washington, DC: NASW Press.
2. Bertrand, M., Mullainathan, S. 2003. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. NBER Working Paper No. 9873; .
3. *Central Hudson Gas & Electric Corp. v. Public Service Commission of New York*. 1980. Supreme Court of the United States, 447 U.S. 557.
4. Dwork, C., Hardt, M., Pitassi, T., Reingold, O., Zemel, R. 2011. Fairness through awareness. arXiv:1104.3913 [cs.CC]; <http://arxiv.org/abs/1104.3913>.
5. Equal Employment Opportunity Commission. 2012. Consideration of arrest and conviction records in employment decisions under Title VII of the Civil Rights Act of 1964. Washington, DC. 915.002; http://www.eeoc.gov/laws/guidance/arrest_conviction.cfm.
6. Equal Employment Opportunity Commission. 1978. Uniform guidelines on employee selection procedures. Washington, DC.
7. Fryer, R., Levitt, S. 2004. The causes and consequences of distinctively black names. *The Quarterly Journal of Economics* 59(3); <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., Keller, K. 2005. Ex-offenders need not apply: the criminal background check in hiring decisions. *Journal of Contemporary Criminal Justice* 21(1): 6-30.
12. Panel on Methods for Assessing Discrimination, National Research Council. 2004. Measuring racial discrimination. Washington, DC: National Academy Press.
13. Pang, B., Lee, L. 2004. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*.
14. Schneider, J. <http://www.lehigh.edu/bio/jill.html> (Archived at <http://foreverdata.org/onlineads>).
15. Sweeney, L. 2013. Discrimination in online ad delivery. (For detailed results and analysis, see full technical report archived 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. 1970. Racism in America and how to combat it. Washington, DC.
17. WebShot Command Line Server Edition. Version 1.9.1.1; <http://www.websitescreenshots.com/>.

LOVE IT, HATE IT? LET US KNOW

feedback@queue.acm.org

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. She earned her Ph.D. in computer science from MIT in 2001. More information about her is available at latanyasweeney.org.

© 2013 ACM 1542-7730/13/0300 \$10.00