robinson +yu

1200 18th St., NW # 807
Washington, DC 2036
+1 (202) 657-9892
http://www.robinsonyu.com

Friday, October 31, 2014

Chairwoman Edith Ramirez
Federal Trade Commission
Room 438
600 Pennsylvania Avenue, NW
Washington, DC 20580

Dear Chairwoman Ramirez:

I was honored to serve as a panelist at the FTC's recent workshop, “Big Data: A Tool for Inclusion or Exclusion?” I now write to provide a copy of a new report from our firm, **Civil Rights, Big Data, and Our Algorithmic Future**, which was discussed during the panel.

Earlier this year, our firm helped coordinate the development of Civil Rights Principles for the Era of Big Data, which were endorsed by leading national civil rights, civil liberties, and media policy organizations.[1] Subsequently, a White House policy review led by John Podesta found that "big data analytics have the potential to eclipse longstanding civil rights protections in how personal information is used in housing, credit, employment, health, education, and the marketplace."[2]

Since then, there has been an outpouring of interest from policymakers, community advocates, corporate leaders and the public. People want to know more about the concrete examples that motivate this work. How and where, exactly, does big data become a civil rights issue? This report begins to answer that question, highlighting key instances where big data and civil rights intersect.

We hope this report will serve as a valuable resource to everyone involved in this important, emerging conversation.

Sincerely,

David Robinson
Principal
Robinson + Yu

---

[1] *Civil Rights Principles for the Era of Big Data*, The Leadership Conference on Civil and Human Rights, http://www.civilrights.org/press/2014/civil-rights-principles-big-data.html (last visited October 29, 2014).
[2] *Big Data: Seizing Opportunities, Preserving Values*, Executive Office of the President (May 2014), http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf.

Civil Rights, Big Data, and Our Algorithmic Future

A September 2014 report on social justice and technology by Robinson + Yu

robinson +yu

Table of Contents

Foreword

As this report goes to print, a tragedy is unfolding in Ferguson, Missouri, where Michael Brown, an unarmed African-American teenager was fatally shot by Darren Wilson, a White police officer. In the months to come, federal and local investigations will seek to determine what happened during the encounter between police officer and civilian.

Consider how differently these investigations might have gone had Officer Wilson used a body-worn camera to record his interaction with Mr. Brown. With strict measures to ensure proper protocols are in place, such cameras can be a powerful tool for police oversight and accountability, as well as to address longstanding deficiencies in police practice that disproportionately impact communities of color. The police in Ferguson have now rolled out such cameras, and a growing number of departments around the country are doing the same. These changes come too late for Mr. Brown, but they will help to make police more accountable for their conduct going forward.

Wade Henderson, President & CEO, The Leadership Conference on Civil and Human Rights

You might call this a big data issue. Or you might say it's about criminal justice reform.

Whether we use the language of big data or civil rights, we're looking at many of the same questions. And that's why this report is so important.

Big data can and should bring greater safety, economic opportunity, and convenience to all people. At their best, new data-driven tools can strengthen the values of equal opportunity and equal justice. They can shed light on inequality and discrimination, and bring more clarity and objectivity to the important decisions that shape people's lives.

But we also see some risks. For example, inaccuracies in databases can cause serious civil rights harms. The E-Verify program, the voluntary, government-run system that employers can use to check whether new employees are work-eligible, has been plagued by an error rate that is 20 times higher for foreign-born workers than for those born in the United States. E-Verify has been under development since it was first authorized in 1996, uses data only from one fairly homogenous source—the government—and is frequently audited. Yet after nearly 20 years, persistent errors remain. This experience provides an important lesson for existing commercial systems, which are fairly new and untested, use data from widely different sources, and operate with no transparency.

Foreword

In February 2014, The Leadership Conference on Civil and Human Rights joined other civil rights and media reform organizations in endorsing the Civil Rights Principles for the Era of Big Data. These principles represent the first time that national civil and human rights organizations have spoken publicly about the importance of privacy and big data for communities of color, women, and other historically disadvantaged groups.

Through these principles, we and the other signatory organizations highlight the growing need to protect and strengthen key civil rights protections in the face of technological change. We call for an end to high-tech profiling; urge greater scrutiny of the computerized decisionmaking that shapes opportunities for employment, health, education, and credit; underline the continued importance of constitutional principles of privacy and free association, especially for communities of color; call for greater individual control over personal information; and emphasize the need to protect people, especially disadvantaged groups, from the documented real-world harms that follow from inaccurate data.

In the coming years, the use of data will have a greater and greater impact on the lives of all people in the United States. To ensure that big data serves the best interests of each of us, civil rights must be a key part of any public policy framework. This report is a critical tool for ensuring that the voices of the civil and human rights community are heard in this important, ongoing national conversation.

Wade Henderson, President & CEO, The Leadership Conference on Civil and Human Rights

Introduction

The key decisions that shape people's lives—decisions about jobs, healthcare, housing, education, criminal justice and other key areas—are, more and more often, being made automatically by computers. As a result, a growing number of important conversations about civil rights, which focus on how these decisions are made, are also becoming discussions about how computer systems work.

Earlier this year, a path-breaking coalition of major civil rights and media justice organizations released the **Civil Rights Principles for the Era of Big Data**, highlighting how the growing use of digital surveillance, predictive analytics, and automated decision-making impacts core civil rights concerns. We served as technical advisors to that coalition.

After the release of the Principles, there was an outpouring of interest from policymakers, community advocates, corporate leaders and the public. People want to know more about the concrete examples that motivate this work. How and where, exactly, does big data become a civil rights issue? This report begins to answer that question, highlighting key instances where big data and civil rights intersect. We hope it will serve as a valuable resource to everyone involved in this important, emerging conversation.

— David Robinson, **Harlan Yu**, and **Aaron Rieke**, **Robinson + Yu**

Chapter 1: Financial Inclusion

As Insurers Embrace Big Data, Fewer Risks Are Shared

Since 2011, Progressive has offered Snapshot, a small monitoring device that drivers must install in their cars to receive the company's best rates.[1] The company offers discounts when the device reports that a driver brakes smoothly, keeps off the roads late at night, and drives infrequently—behaviors that correlate with a lower risk of future accidents.

Low-income individuals, many of whom are people of color, are more likely to work the night shift, putting them on the road late at night, and to live further from work.[2] Devices like Snapshot reduce rates for some drivers by reducing the overall amount of risk sharing among drivers on the road, which means relatively higher costs for those with long car commutes or graveyard shift jobs. At the same time, such systems put responsible late shift workers into the same small category with late-night party-goers, forcing them to carry more of the cost of intoxicated and other irresponsible driving that happens disproportionately at night. Statistically speaking, this added cost does not simply reflect the risk that the late night commuter may be hit by a drunk driver. It also reflects the possibility that, as far as the insurer can tell, the late responsible night worker may be a drunk driver.

> "Big data" allows for a new level of specificity in underwriting, changing how risk is allocated.

Insurers and lenders have long relied on statistics to help them assess the risks of prospective customers. But the deluge of "big data" allows for a new level of specificity in underwriting, changing how risk is allocated. Spreading risk among the insured population is a fundamental purpose of insurance. Some forms of price differentiation, such as charging more to drivers who accelerate or brake suddenly, may provide valuable incentives for the insured to drive more carefully—incentives to which drivers can respond by changing the way they drive. But for people who have to drive at night in order to reach their jobs, this differential pricing provides no benefit. It is simply an added cost.

A person's future health, like their driving behavior, can also be predicted based on personal tracking to set insurance prices. At an annual conference of actuaries, consultants from Deloitte explained that they can now use thousands of "non-traditional" third party data sources, such as consumer buying history, to predict a life insurance applicant's health status with an accuracy comparable to a medical exam.[3] Models based on these data can "predict if individuals are afflicted with any of 17 diseases (e.g. diabetes, female cancer, tobacco related cancer, cardiovascular, depression, etc.) which impact mortality." Deloitte's model also incorporates the health of an applicant's neighbors, at scales as small as two city blocks.

More individualized insurance pricing promises lower rates for those with the lowest risk. At the same time, however, this underwriting means less sharing of risk. Healthy people in low-income

neighborhoods will pay more for their life insurance than will healthy people in healthier neighborhoods (because they are saddled with the health costs of their less healthy neighbors).[4] Responsible night drivers will pay more for car insurance than will responsible daytime drivers (reflecting not only the night driver's risk of being hit by a drunk driver, but also the risk that, as far as the insurer knows, the night driver might be a drunk driver). Insurance prices that are more accurate for most people may, by the same token, be less fair to those nearest the most vulnerable.

Data Brokers Enable Targeting of Financially Vulnerable Communities

Both the Federal Trade Commission and the Senate Commerce Committee recently released significant research reports on the data broker industry, which collects enormous volumes of information on hundreds of millions of Americans. The reports detail how these largely-unregulated companies enable precision-marketing of consumer products to financially vulnerable individuals. The Senate report further warned that the data sold by some brokers is "likely to appeal to companies that sell-high cost loans and other financially risky products," and the FTC observed that many would find it "disconcerting," to know that products can easily be targeted at disadvantaged people. [5]

The lists enable marketers to identify vulnerable consumers with ease.

The Senate report identified marketing lists with titles like "'Rural and Barely Making It,' 'Ethnic Second-City Strugglers,' 'Retiring on Empty: Singles,' 'Tough Start: Young Single Parents,' and 'Credit Crunched: City Families.'" [6] The Commission's report also highlighted segments focused on minority communities and low-income individuals, including a one called the "Urban Scramble."[7] It also observed that data brokers sell "Assimilation Codes," indicating a person's degree of familiarity with the English language. [8] Much of the negative publicity these marketing lists have received stems from their evocative titles—but the fundamental issue runs deeper: the lists enable marketers to identify vulnerable consumers with ease.

Of course, targeted marketing has a place in connecting all communities with the products and services most attractive to them—including for poor consumers, people of color, and people who speak different languages. But precision targeting of vulnerable groups also carries a risk of harm.

Modern data brokerage is an evolution of an old practice. Businesses have a long history of collecting data to help them target or acquire new customers. However, information technology has facilitated a rapid increase in both the volume and availability of data about individuals. Companies are now able to collect and store far more than would have been thought possible in decades past.

"Data broker" is a broad label used to describe the companies that buy, sell, or analyze consumer information. These firms offer marketing services, fraud prevention, risk assessment,data consolidation, or just resell data to other data brokers. There is no comprehensive list of companies that fall under this umbrella. [9]

Data brokers vacuum up data from wherever they can, including from public records, social media sites, online tracking, and retail loyalty card programs. Using these data, brokers build "modeled" profiles about individuals, which include inferences and predictions about them. For example, a broker might infer marital status from the prefix "Mrs." or wealth based on an individual's neighborhood. These profiles are often sold in the form of "segments" (or marketing lists) which are priced and sold by the thousands.

There are few laws governing the data brokerage industry.

There are few laws governing the data brokerage industry, even though many of its practices can resemble the type of consumer scoring that is regulated other contexts. The Government Accountability Office explained that "consumers generally do not have the right to control what personal information is collected, maintained, used, and shared about them—even where such information concerns personal or sensitive matters about an individual's physical and mental health." [10] Similarly, federal law gives consumers the right to correct errors in their credit histories, but no similar right exists with respect to the profiles held by data brokers. The data brokerage industry has been repeatedly criticized for its lack of transparency, and the FTC recently unanimously renewed its call for Congress to enact legislation and empower individuals by allowing them access to information held by data brokers. [11]

This unregulated landscape is a challenge to social justice groups who are mindful of a history of predatory marketing and lending toward vulnerable groups. Data brokers can enable discriminatory targeting based on sensitive information like financial situation, health indicators, or other signs of vulnerability.

Furthering Financial Inclusion with "Alternative Data"

A lack of high-quality, individualized financial data can exclude a person from the mainstream financial system.

Credit is often extended on the basis of an individual's credit score. [12] Today, most credit scores are generated from credit reports, which are maintained by national credit bureaus. Credit reports contain a somewhat limited set of financial indicators, including data about existing credit cards and loans. Traditional credit scores have been shown to be accurate in predicting consumers' creditworthiness (that is, the chance that the consumer will repay credit in accordance with its terms). [13] But not all individuals have a credit report with enough data to generate a credit score. Thus, in some cases, a lack of high-quality, individualized financial data can exclude an individual from the mainstream financial system.

According to the National Credit Reporting Association, as many as 70 million Americans do not have a credit score, or have a lower score than their full financial history would warrant. [14] Because many of these so-called "no file" or "thin-file" individuals regularly pay their utility and phone bills, some groups have argued that this payment data (which is currently not included in most credit files) should be routinely reported to credit bureaus. The major credit bureaus agree, and have developed scoring algorithms that can consider this so-called "alternative" data when it is included in a credit report. [15]

The industry-funded Policy and Economic Research Council (PERC) claims that there is "overwhelming and incontrovertible" evidence that including bill repayment data in credit scores would help low-income individuals. [16] It argues that most people will benefit when such data is included, particularly low-income individuals. This is true, the group continues, "whether the metric is credit score changes, credit score tier changes, or changes in portfolio acceptance given a target default rate." [17] PERC uses these arguments to urge advocates to make the financial system "more inclusive by making credit files more inclusive." [18]

But the National Consumer Law Center (NCLC) has arrived at different conclusions. It claims that the industry is motivated in part by a desire to force utility bills to the "top of [consumers'] payment pile," where such bills might go if they became a factor in access to credit. [19] It also emphasizes that if short-term delinquent payments become part of a credit file, "many low-income customers would receive negative credit reporting marks." [20] Finally, it worries that reporting of utility payments would conflict with established state regulatory policies designed to protect low-income individuals, who may "sometimes defer full payment of utility bills, knowing they are protected from shutoff." [21] In short, concluded NCLC, "[f]ull utility credit reporting will cause disproportionate harm to low-income consumers." [22]

Complicating matters, credit reports are also used to evaluate individuals for jobs, screen applicants for apartment rentals, and generate "marketing scores" for use in marketing consumer products. The impacts of these uses have not been tested or evaluated with the same rigor or transparency as the central use case of consumer credit underwriting, and there are risks that such non-credit uses of credit scores may have a disproportionate adverse impact on protected status groups. Some

protections are in place: for example, most states now have some rules in place to regulate the use of credit information for insurance underwriting. [23] But as the use of credit data continues to expand, so too must the regulatory scrutiny as to the accuracy, fairness, and aggregate impact of such uses. Even if new data would be helpful in the specific context of credit, a broader debate that encompasses the other regulated uses of credit scores is needed.

Alternative data represents both an opportunity and a challenge for the civil rights community.

Alternative data represents both an opportunity and a challenge for the civil rights community. There is some strong evidence suggesting that alternative data could benefit marginalized groups, but data proving this argument has not yet been made available to the civil rights community. To date, stakeholders have not been given the opportunity to reproduce the studies published by industry groups like PERC—much of the underlying data remains proprietary. Greater transparency regarding the impacts of including new data have important work to do in making sure that none of these changes harm vulnerable groups. [24]

Chapter 2: Jobs

E-Verify: The Disparate Impact of Automated Matching Programs

E-Verify is an online database run by U.S. Citizenship and Immigration Services (USCIS), a component of the Department of Homeland Security (DHS). [25] It is designed to help employers quickly determine whether or not newly hired workers are legally eligible to work in the United States. Unfortunately, systematic problems have caused many eligible workers to lose their jobs or to face pre-employment discrimination. Recent studies have shown that these burdens fall disproportionately, and sometimes illegally, on minority groups, including lawful permanent residents and other authorized immigrants.

Today, more than 500,000 U.S. employers use E-Verify when hiring new workers. [26] When a new worker is hired, the employer enters the information from the new hire's I-9 form into the E-Verify website. The site compares the worker's information against multiple government databases held by DHS and the Social Security Administration (SSA). [27] These databases increasingly include data pulled from other local, state and federal agencies. E-Verify informs the employer that the new hire is work eligible, or else produces a Temporary Non-Confirmation (TNC) that the worker may not be eligible to work in the United States. Unfortunately, the process for contesting a database error is expensive and time consuming. And employers may not have the patience to deal with it.

While E-Verify's matching process seems straightforward, many technical and operational issues contribute to erroneous determinations, particularly for noncitizens.

Despite their importance, the algorithms used by E-Verify are not disclosed to the public.

For example, long names are often truncated in USCIS and SSA databases or on printed documents (and truncated different ways in different places) which leads to confusion for employers and, ultimately, mismatches in the system. [28] Matching algorithms determine how strict or lenient the matching process will be. A matching algorithm could allow the first name and the last name to be swapped (e.g., to account for cultures where the family name is printed first rather than second) or could ignore missing punctuation (e.g., to account for cultures with higher frequencies of hyphenated or typographically complex names). Other algorithms could be stricter, requiring an exact character-by-character match. In practice, these algorithms are more complicated, and may use a combination of strategies to tune their accuracy levels. But despite their importance, the algorithms used by in the E-Verify process are not disclosed to the public. [29]

Another common kind of mismatch occurs when a worker changes his or her name—say, because of a recent marriage—and the database still contains that person's prior or maiden name. Until the database is updated, it will retain stale information about the worker. The matching algorithm may

determine that the worker isn't authorized because the system does not know that the new, married name corresponds to an authorized person. In one set of cases where E-Verify errors were successfully resolved, 94 percent of the errors were traced to the worker's having legally changed his or her name. [30]

When the system cannot verify the worker's eligibility, it issues a TNC to the employer, who is required to notify the employee about the adverse determination. If the employee believes that the TNC is a mistake, he must then contest the determination with the government—a painstaking process that can take weeks.

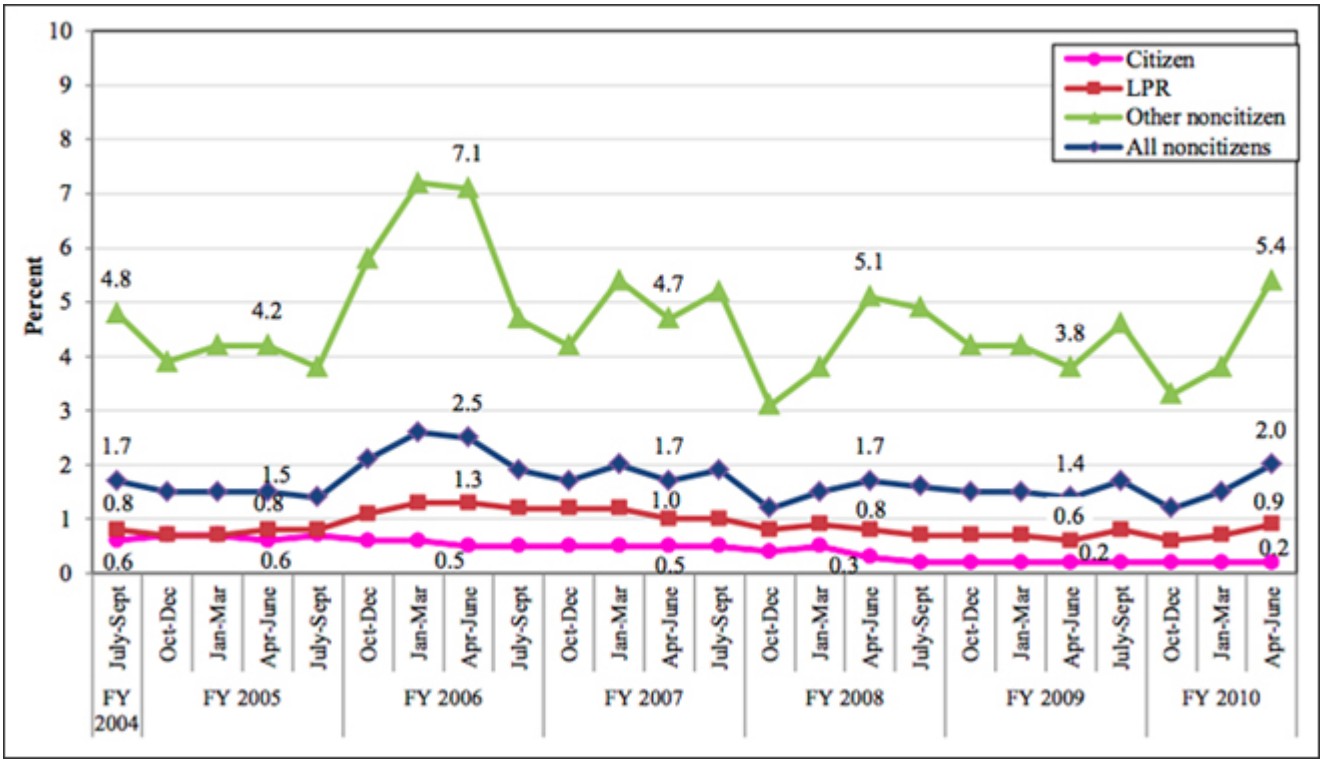

Exhibit III-4. Erroneous TNC Rates, by Attested Citizenship Status: July 2004-June 2010

A recent DHS-funded study found a major discrepancy between the erroneous TNC rates for citizens and noncitizens. [31] It found that legal permanent residents (LPRs) were nearly five times more likely than citizens to be issued an inaccurate TNC, even though they were employment authorized (0.9% for LPRs versus 0.2% for citizens). That figure is even worse for other noncitizens, which were twenty-seven times more likely to receive an inaccurate TNC (5.4%).

Employers have restricted work, delayed training, reduced pay, and taken other unlawful actions against workers who receive TNCs. [32] Because of the uncertainty caused by TNCs, the National Immigration Law Center suggests that E-Verify "encourages employers to hire U.S. citizens exclusively, a practice that usually constitutes a violation of antidiscrimination law." [33]

E-Verify is voluntary for most employers today [34] but there have been recent legislative efforts to make the program mandatory nationwide. [35] While automated technologies can bring vast

efficiency improvements to many government processes, including this one, even small error rates that result from seemingly minor technical issues can have a life-changing impact on thousands of people.

E-Verify's matching and verification problems are far from unique. Inaccurate information in databases, and the inability to consistently fix errors, have also been a major pain point in a number of other areas. The credit reporting industry has offered an Orwellian struggle to people with errors in their credit reports [36] (a still-bad situation that may be slowly improving [37]); a voter ID law in Texas requiring an exact match between the state's voting rolls and the name on the voter's identification card has caused difficulties for thousands of married women at the polls, forcing them to sign affidavits and file provisional ballots. [38] Error rates that look small on a spreadsheet can loom large in the lives of the people affected.

Hiring Algorithms May Put Jobs Out of Reach

Many retailers, call centers, and other employers of entry-level service staff have begun using machine learning systems to evaluate job applicants. Analyzing numerous factors for thousands of employees, specialized technology firms develop online questionnaires that surface the factors most predictive of success for each employer and job.

Some firms have found that people with shorter commutes tend to make better hires, because they are statistically likely to stay in the job longer. This insight may be particularly important for service sector employers, whose hiring is increasingly automated, and for whom turnover is a major concern. According to a 2012 Wall Street Journal report, a hiring analytics firm called Kenexa (now owned by IBM) "asks applicants for call-center and fast-food jobs to describe their commute by picking options ranging from 'less than 10 minutes' to 'more than 45 minutes.' The longer the commute, the lower their recommendation score for these jobs, says Jeff Weekley, who oversees the assessments." [39] The same story also notes that how reliable a person's transportation is (i.e., whether they depend on public transportation) and how long they have lived at their current address may also be considered.

A second firm that applies big data to the hiring process, Evolv, has reportedly made a different choice. As the Atlantic Monthly reported:

There are some data that Evolv simply won't use, out of a concern that the information might lead to systematic bias against whole classes of people. The distance an employee lives from work, for instance, is never factored into the score given each applicant, although it is reported to some clients. That's because different neighborhoods and towns can have different racial profiles, which means that scoring distance from work could violate equal-employment-opportunity standards. [40]

A hiring preference against workers who live far away may be accurate—they may really average shorter tenure in the job—but is it fair?

A hiring preference against workers who live far away may be accurate—they may really average shorter tenure in the job—but is it fair? Such a preference punishes people for living far from where the jobs are, and can particularly hurt those living in economically disadvantaged areas, who are disproportionately people of color. Such practices make it even harder for people in disadvantaged communities to work their way out of poverty.

In Online Searches, Big Data Systems Reproduce Racial Bias

Digital indicators of race, religion, or sexual preference can easily be observed or inferred online. In some ways, these indicators are just like those an employer might pick up when scanning a person's resume. [41] However, a recent study has revealed that these indicators can foster "discriminatory outcomes or giv[ing] preference to members of one group over another" when combined with complex big data systems. [42]

> There is discrimination in delivery of these ads.

Latanya Sweeney, a computer science professor at Harvard who recently served as Chief Technologist at the Federal Trade Commission, described how Google ads discriminate based on the name of the person searched. [43] When searching for her own name on Google, Dr. Sweeney noticed ads referencing arrest records. This prompted her to design a study to learn whether searches for white-identifying names prompted the same sorts of ads as searches for black-identifying names did. She found that a greater percentage ads with "arrest" in their text appeared for black-identifying names than for white-identifying names, to an extent that could not plausibly be explained by chance. [44] She concluded that "[t]here is discrimination in delivery of these ads."[45]

This happens because Google's software automatically learns which ad combinations are most effective (and most profitable) by tracking how often users click on each ad. These user behaviors, in aggregate, reflect the biases that currently exist across society. Instantcheckmate.com, a leading company that sells arrest records, denied that it has ever tried to connect a name with race. But it would not necessarily have to for this outcome to occur. [46]

Ads that are more often clicked on automatically receive a higher "quality score"—and are more often displayed—in Google's system. [47] Google and InstantCheckmate may automatically find themselves reinforcing the racial biases that their audience's click patterns reflect. Dr. Sweeney explains: "If Google's Adsense service learns which ad combinations are more effective, it would first serve the arrest-related ads to all names at random. But this would change" as the algorithm automatically changed in response to a pattern, where "click-throughs are more likely when these ads are served against a black-identifying name." [48]

These sorts of structural discrimination issues are particularly troubling as employers—and others in positions of power and responsibility—increasingly consult the Internet when making the decisions that shape people's lives. [49] Although potential employees have some legal protections today, it would be difficult for a job applicant harmed by the subliminal effects of biased ads to trace such harm to its cause. A quick glance (or many such glances) by a hiring professional are likely to go unnoticed. The same concerns may arise in situations involving promotions, special awards, or other forms of professional advancement, or in different settings such as the search for a roommate.

Lawyers do caution employers to tread carefully online. "I advise employers that it's not a good idea to use social media as a screening tool," says James McDonald, a specialist in employment law. [50] "[Employers] need to control the information," he says, but the ease of a Google search may be hard

to resist. “By and large, employers avoid asking questions about these traits in interviews. But now technology makes it easier to find that information,” observes Prof. Alessandro Acquisti of Carnegie Mellon University. [51]

Dr. Sweeney’s research shows that racism can be perpetuated inadvertently by complex online systems, even when the companies that create these systems do not intend to discriminate.

Chapter 3: Criminal Justice

Predictive Policing: From Neighborhoods to Individuals

There is no public, comprehensive description of the algorithm's input.

In February 2014, the Chicago Police Department (CPD) made national headlines for sending its officers to make personal visits to residents considered most likely to be involved in a violent crime. The selected individuals were not necessarily under investigation, but had histories that implied that they were among the city's residents most likely to be either a victim or perpetrator of violence. The officers' visits were guided in part by a computer-generated "Heat List": the result of an algorithm that attempts to predict involvement in violent crime. City officials have described some of the inputs used in this calculation—it includes some types of arrest records, for example—but there is no public, comprehensive description of the algorithm's input.

The visits were part of a new "Custom Notification Program," which sends police (or sometimes mails letters) to peoples' homes to offer social services and a tailored warning.[52] For example, officers might offer information about a job training program or inform a person that federal law provides heightened sentences for people with certain prior felonies.[53] The city reports that the contents of a notification letter are based on an analysis of "prior arrests, impact of known associates, and potential sentencing outcomes for future criminal acts."[54] Although some of these visits have been poorly received,[55] the department argues that the outreach efforts may already have deterred crime.[56] Mayor Emanuel recently claimed that, of the 60 interventions that have already taken place, "none of the notified individuals have been involved in any new felony arrests."[57]

The Heat List is a rank-order list of people judged most likely to be involved in a violent crime, and is among the factors used to single people out for these new notifications. The CPD reports that the heat list is "based on empirical data compared with known associates of the identified person."[58] However, little is known about what factors put people on the heat list, and a FOIA request to see the names on the list was denied on the grounds that the information could "endanger the life or physical safety of law enforcement personnel or [some] other person."[59] Media outlets have reported that various types of data are used to generate the list, including arrests, warrants, parole status, weapons and drug-related charges, acquaintances' records, having been a victim of a shooting or having known a victim,[60] prison records, open court cases, and victims' social networks.[61] The program's designer, Illinois Institute of Technology (IIT) Professor Miles Wernick, has denied that the "algorithm uses 'any racial, neighborhood, or other such information' in compiling the list."[62]

Cities across the country are expanding their use of data in law enforcement. The most common applications of predictive technology are to assist in parole board decisions[63] and to create heat maps of the most likely locations of future criminal activity in order to more effectively distribute

police manpower. Such systems have proven highly effective in reducing crime, but they may also create an echo chamber effect as crimes in heavily policed areas are more likely to be detected than the same offenses committed elsewhere. This effect may lead to statistics that overstate the concentration of crime, which can in turn bias allocations of future resources.

Chicago's experiment is one of several of a new type, in which police departments move beyond traditional geographic "crime mapping" to instead map the relationships among city residents. Specifically, identifying individuals for tailored intervention is the trend most likely to expand in the future of predictive policing—raising important questions on how to ensure justice continues to be protected through machine systems. Other districts are already working with academics to develop similarly styled programs, including one in Maryland that aims to "predict which of the families known to social services are likely to inflict the worst abuses on their children."[64] In projects like these, automated predictions of future bad behavior may arise—and may be acted upon—even without direct evidence of wrongdoing. Such systems will sometimes make inaccurate predictions, and when they do, their mistakes may create unjustified guilt-by-association, which has historically been anathema to our justice system.

Even as they expand their efforts to collect data, city governments often do not have the academic resources to analyze the vast amounts of data they are aggregating. They are often partnering with private or academic institutions to assist in the process. In Chicago, the city is working with the MacArthur-backed Crime Lab to analyze the effectiveness of various programs, including things like "Becoming A Man," a program that focuses on violence prevention among at-risk youth.[65] These partnerships allow the city to expand the ways it uses the data it collects, and may unlock significant benefits (by, for example, demonstrating the effectiveness of non-punitive crime reduction programs). At the same time, the private actors conducting these and other analyses should be held to at least the same standards of accountability and transparency that would apply if the city were analyzing its data internally.

Secrecy Is a Barrier to Responsible Use of Police Technologies

Police departments across the country are adopting new surveillance technologies, including cameras that scan license plates, tablets that recognize faces, and devices that intercept signals from mobile phones. Unfortunately, these tools are often shrouded in secrecy. Police departments deploy new technologies without clear use policies to guide their use or audit logs to keep users accountable,[66] sign non-disclosure agreements (NDAs) with technology vendors,[67] decline public records (FOIA) requests,[68] and cooperate with federal officials to shield key information from disclosure.[69]

If adopted and deployed in an opaque fashion, new police tools open the door for discrimination and other abuses.

Transparency is vital. It allows for the incorporation of safeguards and policies that could pave the way for fair uses of these new technologies. If adopted and deployed in an opaque fashion, new police tools open the door for discrimination and other abuses.

Recent deployments of automatic license plate readers (ALPRs) illustrate this point. ALPRs are high-speed cameras that photograph all passing vehicles, feeding a database that may eventually describe many peoples' comings and goings. Unfortunately, the devices have been used to target minority groups. For example, police officers in New York reportedly drove "unmarked vehicles equipped with license plate readers around local mosques in order to record each attendee."[70] And according to one New York police officer, the use of license plate readers "is only limited by the officer's imagination."[71]

These reports are troubling, but it's easy to imagine ALPRs responsibly deployed. For example, an ALPR system could discard scans that don't raise red flags (fewer than 1% do). Police departments could easily develop better policies by asking where ALPRs are deployed, when and how data can be used, and how long it is retained. Unfortunately, today, ALPR data is "often retained for years or even indefinitely, with few or no restrictions . . . ."[72]

It can be exceedingly difficult for advocates to learn about police technologies. Police departments across the country sign NDAs from technology vendors, which function as gag orders. For example, an Arizona journalist obtained a copy of the agreement between a vendor called Harris Corporation (which produces "Stingrays," a high-tech device that mimics a cell phone tower in order to snoop on nearby mobile phones) and the Tucson Police Department.[73] Among other things, it barred the police department from discussing the technology with any government entity and required the department to notify Harris whenever it received a public record request concerning one of the company's "protected products."[74] These NDAs are sometimes cited when declining public record requests, even if they seem to conflict with public record laws.[75]

In another case, in the summer of 2014, U.S. Marshals drove to a local police station in Sarasota, Florida, seizing some of the station's Stingray-related records and moving them to an undisclosed location.[76] The ACLU had evidence that police were using the devices without a probable cause warrant, relying instead on a less protective state statute typically used for more limited requests. "We've seen our fair share of federal government attempts to keep records about [cell phone surveillance technology] secret, but we've never seen an actual physical raid on state records in order to keep them from public view," said Nathan Wessler, an attorney for the ACLU.

Advocates aren't the only ones left in the dark: police technologies have also been hidden from the courts. For example, shortly after the seizure of records by U.S. Marshals, the ACLU released email messages showing that Florida police repeatedly and deliberately hid their use of cell phone surveillance technology from state courts.[77] "Concealing the use of stingrays deprives defendants of their right to challenge unconstitutional surveillance and keeps the public in the dark about invasive monitoring by local police," noted Maria Kayanan, Associate Legal Director of the ACLU of Florida.

These problems are exacerbated by federal grant programs—including those that give police access to decommissioned military hardware, and the "equitable sharing" regime for federally seized drug funds. Such programs allow police to sidestep the natural oversight that happens when new police investments rely on local funds. This culture of secrecy threatens civil rights and short-circuits a needed debate about how to use these tools responsibly.

There are many opportunities for the civil rights community to help create appropriate boundaries for use of new police technologies. Some positive steps have already been taken, such as the 2007 DHS Best Practices for Government Use of CCTV.[78] Deployment of some police technologies could be contingent on their use being bounded by traditional constitutional safeguards, such as judicial approval and warrants, whether implemented through federal or state law or other means. However, before appropriate debates about the uses of the newest technologies can happen, law enforcement must show a willingness to enter into a dialogue and move their technologies out of the shadows.

Body-Worn Cameras Can Promote Police Accountability

New technology allows every police officer to wear a small camera that makes digital video recordings of citizen interactions.[79] In Ferguson, Missouri, where conflicting accounts of the August 9 shooting of unarmed 18-year-old Michael Brown led to rioting and worldwide news, the police had received a stock of such cameras, but they had not yet deployed when Brown was shot.[80] If body worn cameras had been rolling, they might have provided a key record of the fateful interaction—or even, by their presence, encouraged more cautious behavior and prevented the incident from occurring at all. In the wake of Brown's shooting and a string of other deaths of unarmed black men, a number of civil rights groups including the NAACP Legal Defense & Education Fund have urged wider use of the cameras.[81] They have also been proposed as a remedy to New York City's stop and frisk practices; mayor de Blasio called for the cameras' use as part of his campaign,[82] and the city will soon begin testing use of the cameras.[83]

According to a report by the Police Executive Research Forum, as of September 2013, one out of out of every four police departments it surveyed said that they had "begun to equip their officers with body-worn cameras on at least a trial basis."[84] More recently, a researcher working for the group reported that Los Angeles, London, and a number of other large cities are piloting the cameras.[85] There are many different types of body-worn cameras available, including some that mount to vests or headgear. They may cost as much as $1,000 each, or as little as a few hundred dollars, depending on the model.[86]

The impact of body-worn cameras has not yet been systematically studied.[87] "The technology is ahead of the policy at this point," noted James Stewart, director of public safety and security for CNA Analysis and Solutions.[88] "[N]obody has done a real, large-scale research study on the effect of cameras on whether it reduces injuries, complaints, and whether the people wearing them feel comfortable wearing them."

But anecdotal evidence is generally positive, suggesting that the cameras can help de-escalate heated situations.[89] For example, a small police department in Rialto, California equipped half of its 54 uniformed officers with cameras, after which complains against police dropped 88 percent compared with the previous 12 months.[90] And in Oakland, California, police have deployed around 500 cameras (one of the largest programs in the nation).[91] "It's enormously helpful," reflected Oakland Police Interim Assistant Chief Paul Figueroa. "When you're able to go to the video and see what's occurred . . . it saves so much investigative time, all the way around as to whether the misconduct took place or not."[92]

The idea of body-worn cameras is even spreading at the federal level. In March, Secretary of Homeland Security Jeh Johnson met with immigrant rights activist groups to discuss the possibility of border patrol agents wearing cameras.[93] The Obama Administration is hoping that the recordings will provide insight into complaints of excessive force on the border.

Civil liberties groups generally support the cameras, but urge some caution. "For the ACLU, the challenge of on-officer cameras is the tension between their potential to invade privacy and their strong benefit in promoting police accountability," writes Jay Stanley of the American Civil Liberties Union.[94] "Overall, we think they can be a win-win—but *only* if they are deployed within a framework

of strong policies." Key issues include whether the footage will become a public record,[95] how the system tracks officer decisions to turn the cameras off or delete footage, and whether citizens can veto recordings in private settings such as their homes.

Body-worn cameras are poised to help boost accountability for law enforcement and citizens.

Body-worn cameras are poised to help boost accountability for law enforcement and citizens. And unlike many new police technologies, the cameras share preliminary support from *both* law enforcement and social justice groups. The cameras may be particularly beneficial for Black and Latino men, who are likelier to shoulder the harms of any police misconduct due to their greater involvement with the criminal justice system. But successful implementation of the cameras will require careful policies that respect and protect both the police and the public.

Chapter 4: Government Data Collection and Use

Dragnet Surveillance Short-Circuits Court Review in Drug Cases

> Secret tools first justified on national security grounds are also being applied for domestic law enforcement purposes.

The federal government's large-scale warrantless surveillance of Americans, first brought to light by Edward Snowden, initially appeared to be a story about national security. But the public has now learned that some of the secret tools first justified on national security grounds are also being applied for domestic law enforcement purposes.

For example, the New York Times has reported that the White House Office of National Drug Control Policy, together with the Drug Enforcement Administration (DEA), has secretly paid AT&T to build a massive database of calling records for domestic law enforcement use.[96] AT&T officially controls this database, known as "Hemisphere," but works closely with the government to give it seamless access, even embedding company employees on law enforcement teams. The program has operated since 2007. The information that AT&T gathers is available not only to the federal government but also to local detectives who are working on drug investigations.[97] Unlike more traditional methods, this system lets law enforcement examine everyone's activity—not just people who are under suspicion. The leaked document published by the Times explains that this system can "often identify cell phones the target is using that are unknown to law enforcement."[98]

A document also shows that the White House office involved with the program asked analysts to "never refer to Hemisphere in any official document," but instead to issue a seemingly independent subpoena for records that had been uncovered by Hemisphere. This controversial, secret tactic—called "parallel construction"—allows law enforcement to offer such records as evidence in court, without alerting the court or the defendant to its reliance on this warrantless program.[99] The same tactic is apparently used by the DEA's "Special Operations Division" (SOD), which maintains a partnership with the NSA to receive tips gathered via NSA surveillance activities, outside the framework of domestic law enforcement.[100] Guidance obtained by Reuters "specifically directs agents to omit the SOD's involvement from investigative reports, affidavits, discussions with prosecutors and courtroom testimony. Agents are instructed to then use 'normal investigative techniques to recreate the information provided by SOD.'"[101]

Nancy Gertner, a former federal judge now on the faculty of Harvard Law School, describes such tactics as "phonying up the course of the investigation."[102] She warns that "[w]hen the DEA is concealing what the source of the information is and pretending it came from one place rather than another, there can be no judicial review" of the government's real investigative tactics.

The Temptation to Abuse Government Databases

Governments at every level collect personal data, with varying degrees of scope and detail. Data is necessary for many parts of government to function, from providing social services, to enrolling students in public schools, to issuing passports. The scale of collection is expansive and growing, but it's also difficult to quantify,[103] particularly because private companies often play a central role in helping governments gather and store data.[104] Data needs to be accessible to government employees and contractors for legitimate purposes. But, in far too many instances, the availability of personal data—and government's inability to properly manage and oversee its own staff's activities—has led to abuse.

At a local level, women appear over-represented among the targets of such abuse, especially by police.

At a local level, women appear over-represented among the targets of such abuse, especially by police. Law enforcement officers have been accused of inappropriately using databases to stalk women in Minnesota,[105] Illinois,[106] New Jersey,[107] New York,[108] and North Carolina.[109] One officer went so far as to travel to a woman's vehicle and leave a note for her.[110] Another extensive episode involved an NYPD detective who "hacked into computers to collect private email accounts and passwords of 21 fellow officers and nine other people to snoop on his ex-girlfriend."[111] Officers have also been accused of using databases to look up individuals at the request of women they were trying to impress in Connecticut[112] and Alaska.[113]

The federal government maintains an enormous amount of data, raising similar issues on a national scale. For example, Edward Snowden recently claimed that it was common for analysts at the NSA to share nude photos they find in the course of their duties.[114] A similar story emerged in 2008 when two former NSA analysts came forward to discuss their access to the personal communications of US citizens living abroad. They reported that the analysts would recommend particular calls to each other based on their sexual content and that the list of calls is "stored the way you'd look at songs on your iPod and you could pull up a song on your iPod" using a person's phone number.[115]

In a series of cases commonly referred to as LOVEINT, the NSA confirmed that, in the last decade, there were at least 12 cases of individual analysts using the SIGINT (signals intelligence) system to inappropriately track individuals, as well as two other cases currently under investigation.[116] In most cases, analysts were searching for information on people they knew socially or romantically.[117] The NSA Director of Compliance argued that the number of incidents by the agency was extremely low compared with its overall activities[118] But most of the incidents were self-reported,[119] so its not unreasonable to believe that the actual number of infractions taking place is higher than the reported number.

Collecting, storing and analyzing massive amounts of personal data enhances the government's ability to help and protect ordinary citizens. But the availability of data can pose a temptation too great for a small percentage of those individuals who have access. As governments expand their use of data, there is significant room to improve the policies and practices to further lower the risk that this data is abused.

The Census: Big Data for Civic Welfare

The United States Census is the government's original big data. Since 1790, the Census Bureau has conducted surveys to enumerate the United States population. Required by the Constitution to apportion representation and taxation among the states, the Census Bureau today maintains the most complete and authoritative records of the American population. These records are used for allocating resources and planning for a variety of social and economic programs. The decennial census and the more detailed American Community Survey (the modern version of the census "long form") have become essential for the functioning of government, civil society, and the private sector. Civil rights groups have a strong stake in the accuracy, completeness, and availability of Census Bureau data.

The decennial census is conducted every ten years and counts every individual in the United States. In 2010, the Bureau collected demographic information on each member of the household.[120] Between 1940 and 2000, households were sampled during the decennial census to collect more detailed information, in what was known as the census "long form."[121] In 2005, these more detailed records were separated into the American Community Survey (ACS), which samples households on a continuous basis to provide more up-to-date data for areas as small as census tracts (average population: 4,000). The ACS samples roughly 2.5 percent of U.S. households each year.[122]

In addition to determining Congressional representation, census data is used to allocate resources for social programs, including Medicaid, maternal and child health programs, transit programs, public housing assistance, Community Development Block Grants, Head Start, Title I education funds and grants for special and vocational education.[123] It is also used for planning for hospitals, nursing homes, clinics, and the location of other health services, and drawing school district boundaries.[124]

For Census 2020, the Bureau is researching ways to incorporate new technology to make the census more efficient and less costly.

For Census 2020, the Bureau is researching ways to incorporate new technology to make the census more efficient and less costly. This is, in part, a necessity: the Bureau has been tasked with running the 2020 Census at the same cost as the 2010 Census, after decades of rapid growth in the cost of census-taking.[125] It is considering using governmental administrative records (ARs)—pre-existing government databases from sources such as Internal Revenue Service, Social Security, TANF, SNAP, and Medicare/Medicaid,[126] as well as from state and local governments—to pin down areas of housing change. This will help to eliminate the need to physically canvass every street and address prior to the start of the population count, identify vacant housing units before sending a census taker to knock on doors, and add people to the count who do not self-respond to the census. The Bureau is also looking at new ways to strategically route census takers during non-response follow-up, optimizing their productivity. And it is researching an Internet response option to reduce costs during the self-response phase.

Unfortunately, ARs can both help and hurt the accuracy of the census. Data collected from ARs may lack detail (such as race and ethnicity, especially for subgroups, and household relationships), may

be out of date, or may underrepresent the hardest-to-count populations. Yet, ARs may be suitable as a last resort, when people would otherwise have been missed even after in-person follow-up interviews. In the 1990 Census, for example, the Bureau used parolee and probationer records to add "between 400,000 and 500,000 persons to the final census counts."[127] Many of those added were young, black men, who had been missed. This practice was abandoned in later years because "about half of [people added using parolee and probationer records] were later estimated to have been enumerated erroneously." But technological improvements might increase the utility of ARs in the future.

A secure Internet response option could significantly improve efficiency of the census. Recent testing of the Internet response option suggests it is likely to only marginally increase the total number of self-responders, meaning that most of those who would use the Internet to respond would otherwise have voluntarily responded by mail. But even a small shift will provide the benefit of reducing the cost of mailing out packets to tens of millions of households, which will free up monetary resources that can better be focused elsewhere—like physically canvassing neighborhoods with low Internet adoption rates. However, an Internet response option must be mobile-friendly in order to maximize its reach to certain minorities who over-index on smartphone adoption, such as Hispanics and Asian Americans.

The Census Bureau faces a difficult challenge ahead: saving money while retaining quality and accuracy. Technological advances and new data sources can help, but cost constraints still threaten to make the 2020 Census less complete than its predecessors.

Acknowledgements

This report was prepared and written by Aaron Rieke, David Robinson and Harlan Yu from Robinson + Yu, with the support of Alethea Lange, Erica Portnoy and Joshua Tauberer.

We would like to thank our colleagues for reviewing and providing thoughtful comments our early drafts, including: Danah Boyd at Data & Society; Chris Calabrese at the ACLU; Michael Connor at Open MIC; Malkia Cyril at Center for Media Justice; amalia deloney at Center for Media Justice; Seeta Peña Gangadharan at Open Technology Institute; Alexander Hart at Freedman Consulting; Jason Lagria at Asian Americans Advancing Justice | AAJC; Terri Ann Lowenthal, an independent expert on the census; Joni Lupovitz at Common Sense Media; Terry Ao Minni at Asian Americans Advancing Justice | AAJC; and above all Corrine Yu, The Leadership Conference on Civil and Human Rights.

We would also like to give special thanks to Wade Henderson from The Leadership Conference on Civil and Human Rights for contributing the Foreword to our report.

The report's outreach and rollout was coordinated by Jennifer Calloway, Michael Khoo, and Stephanie Vanegas from Spitfire Strategies; and Scott Simpson from The Leadership Conference on Civil and Human Rights.

The website's design is the work of Joshua Tauberer and Kaptiv8.

Finally, this report would not be possible without the financial support of the Ford Foundation, and the support and encouragement of Jenny Toomey, Lori McGlinchey and Amy Brown. Thank you.

robinson +yu

References

[1] Erik Holm, *Progressive to Offer Data-Driven Rates*, Wall St. J. (2011), **http://online.wsj.com/news/articles/SB10001424052748704433904576212731238464702** (Other major car insurers now offer similar programs).

[2] Maria Enchautegui, *Nonstandard Work Schedules and the Well-being of Low-Income Families*, Urban Institute (2013), **http://www.urban.org/publications/412877.html**.

[3] Alice Kroll & Ernest Testa, *Predictive Modeling for Life Insurance Seminar* (2010), **https://www.soa.org/files/pd/2010-tampa-pred-mod-4.pdf**.

[4] *See* Health Poverty Action, *Key Facts: Poverty and Poor Health*, **http://www.healthpovertyaction.org/policy-and-resources/the-cycle-of-poverty-and-poor-health/the-cycle-of-poverty-and-poor-health1** (Poverty is "inextricably linked" to poor health.).

[5] United States Senate Committee on Commerce, Science, and Transportation, *A Review of the Data Broker Industry: Collection, Use, and Sale of Consumer Data for Marketing Purposes* (2013), **http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=0d2b3642-6221-4888-a631-08f2f255b577**.

[6] *Id.*

[7] Federal Trade Commission, *Data Brokers: A Call for Transparency and Accountability* (2014), **http://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf**.

[8] *Id.*

[9] United States Senate Committee on Commerce, Science, and Transportation, *supra* note 5.

[10] *Id.*

[11] Federal Trade Commission, *supra* note 7.

[12] Research has shown that Credit is "usually necessary to buy a home, build a business, or send your children to college." Ashoka, *Banking The Unbanked: A How-To*, Forbes (2013), **http://www.forbes.com/sites/ashoka/2013/06/14/banking-the-unbanked-a-how-to**.

[13] Board of Governors of the Federal Reserve System, *Report to the Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit* (2007), **http://www.federalreserve.gov/boarddocs/rptcongress/creditscore/creditscore.pdf**.

[14] Arjan Schutte & Rachel Schneider, *The Predictive Value of Alternative Credit Scores* (2007) **http://www.cfsinnovation.com/node/330262?article_id=330262**.

[15] VantageScore Solutions, *VantageScore Consumer Credit Scoring* (2014), http://www.vantagescore.com.

[16] Michael Turner et al., *A New Pathway to Financial Inclusion: Alternative Data, Credit Building, and Responsible Lending in the Wake of the Great Recession* (2012), **http://www.perc.net/wp-content/uploads/2013/09/WEB-file-ADI5-layout1.pdf**.

[17] *Id*.

[18] *Id*.

[19] National Consumer Law Center, Full Utility Credit Reporting: Risks to Low-Income Consumers (2012), **http://www.nclc.org/images/pdf/energy_utility_telecom/consumer_protection_and_regulatory_issues/ib_risks_of_full_utility_credit_reporting_july2012.pdf**.

[20] *Id*.

[21] *Id*.

[22] *Id*.

[23] National Association of Mutual Insurance Companies, *Credit-Based Insurance Scoring: Separating Facts from Fallacies* (Feb. 2010), **http://iiky.org/documents/NAMIC_Policy_Briefing_on_Insurance_Scoring_Feb_2010.pdf**.

[24] Turner et al., *supra* note 16.

[25] U.S. Citizen and Immigration Services, *E-Verify*, **http://www.uscis.gov/e-verify.**

[26] U.S. Citizen and Immigration Services, *Half a Million Companies Now Participate in E-Verify* (Jan. 23, 2014), **http://www.uscis.gov/news/news-releases/half-million-companies-now-participate-e-verify-0**

[27] Westat, *Evaluation of the the Accuracy of E-Verify Findings* 8-10 (2012), **http://www.uscis.gov/sites/default/files/USCIS/Verification/E-Verify/E-Verify_Native_Documents/Everify%20Studies/Evaluation%20of%20the%20Accuracy%20of%20EVerify%20Findings.pdf** (listing the names of no fewer than 13 government databases used by E-Verify).

[28] *Id*. at xv.

[29] *Id*. at 6.

[30] *Id*. at 54-55. ("Names changed through marriage, divorce, during the naturalization process, or otherwise constitute a major reason for name mismatches. SSA Service Representatives reported that most name changes they see are due to marriage and changes of name during the naturalization process.")

[31] *Id*. at 25.

[32] National Immigration Law Center, *Verification Nation* 6 (Aug. 2013), **www.nilc.org/document.html?id=957**.

[33] *Id*.

[34] The federal government requires agencies and federal contractors to use E-Verify. Some states also require businesses to use E-Verify as a condition to apply for and maintain a business license.

[35] *See* Congressional Research Service, *Immigration Legislation and Issues in the 113th Congress* (Nov. 20, 2013), **http://fas.org/sgp/crs/homesec/R43320.pdf**.

[36] *See* Chi Chi Wu, *Automated Injustice* (2009), **http://www.nclc.org/images/pdf/pr-reports/report-automated_injustice.pdf**.

[37] Consumer Financial Protection Bureau, *Now You Have Better Options to Dispute a Credit Report Error* (Feb. 27, 2014), **http://www.consumerfinance.gov/blog/now-you-have-better-options-to-dispute-a-credit-report-error**.

[38] Wade Goodwyn, *Texas Voter ID Law Creates a Problem for Some Women*, NPR (Oct. 30, 2013), **http://www.npr.org/2013/10/30/241891800/texas-voter-id-law-creates-a-problem-for-some-women**.

[39] Joseph Walker, *Meet the New Boss: Big Data*, Wall St. J. (Sep. 20, 2012), **http://online.wsj.com/news/articles/SB10000872396390443890304578006252019616768**.

[40] *See* Don Peck, *They're Watching You At Work*, The Atlantic (Dec. 2013), **http://www.theatlantic.com/magazine/archive/2013/12/theyre-watching-you-at-work/354681**.

[41] David R. Francis, *Employers' Replies to Racial Names*, National Bureau of Economic Research (2003), **http://www.nber.org/digest/sep03/w9873.html**.

[42] Latanya Sweeney, *Discrimination in Online Ad Delivery*, 11 Queue 10 (2013), **http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2208240**.

[43] *Id*.

[44] *Id*.

[45] *Id*.

[46] *Racism is Poisoning Online Ad Delivery, Says Harvard Professor*, MIT Technology Review (2013), **http://www.technologyreview.com/view/510646/racism-is-poisoning-online-ad-delivery-says-harvard-professor**.

[47] See Google, *Check and understand Quality Score*, **https://support.google.com/adwords/answer/2454010** (explaining to advertisers that "your ad's expected CTR [click through rate]" contributes to its Quality Score, which in turn "typically lead[s] to lower costs and better ad positions" for ads with high Quality Scores).

[48] *Racism is Poisoning Online Ad Delivery, Says Harvard Professor*, *supra* note 46.

[49] Nick Fishman, *Survey Shows 48% of Employers Conduct Social Media Background Checks*, IQ Blog (2012), **http://www.employeescreen.com/iqblog/48-of-employers-conduct-social-media-background-checks**.

[50] Jennifer Valentino-Devries, *Bosses May Use Social Media to Discriminate Against Job Seekers*, Wall St. J. (2013), **http://online.wsj.com/news/articles/SB10001424052702303755504579208304255139392**.

[51] *Id*.

[52] Chicago Police Department, *D13-09, Custom Notifications In Chicago – Pilot Program* (2013), **http://directives.chicagopolice.org/directives-mobile/data/a7a57bf0-13fa59ed-26113-fa63-2e1d9a10bb60b9ae.html** ("The letter will be specific to the identified individual and incorporate those factors known about the individual inclusive of prior arrests, impact of known associates, and potential sentencing outcomes for future criminal acts.").

[53] Thomas Frisbie, *Chicago Police 'custom notifications': Is it profiling?*, Chicago Sun-Times (Feb. 26, 2014), **http://voices.suntimes.com/early-and-often/backtalk/chicago-police-custom-notifications-is-it-profiling** ("[F]ederal law says that if you have certain felonies in your criminal history and you get caught with a gun you can be prosecuted as a career armed criminal and [can] face a minimum of 15 years in prison[.]").

[54] Chicago Police Department, *supra* note 52. ("The Custom Notification is predicated upon national research that concluded certain actions and associations within an individual's environment are a precursor to certain outcomes should the individual decide to or continue to engage in criminal behavior.")

[55] Jeremy Gorner, *Chicago Police Use 'Heat List' as Strategy to Prevent Violence,* Chicago Tribune (Aug. 21, 2013), **http://articles.chicagotribune.com/2013-08-21/news/ct-met-heat-list-20130821_1_chicago-police-commander-andrew-papachristos-heat-list** ("'I haven't done nothing that the next kid growing up hadn't done. Smoke weed. Shoot dice. Like seriously?' an incredulous McDaniel said while recalling the recent visit from police brass with a Tribune reporter.").

[56] *Id.*

[57] Robin Kelly, *Kelly Report 2014: Gun Violence in America* 15 (2014), **http://robinkelly.house.gov/sites/robinkelly.house.gov/files/wysiwyg_uploaded/KellyReport_1.pdf**.

[58] Chicago Police Department, *supra* note 52.

[59] Letter from P.O. Cronin, Assistant Freedom of Information Officer, Chicago Police Dep't, to Matthew Stroud, Reporter, The Verge (Jan. 6, 2014), **http://cdn2.sbnation.com/assets/4020793/Stroud-CPD-FOIA.jpg** ("Endanger the life or safety of law enforcement personnel or any other person.").

[60] Kristal Hawkins, *Heat list' brings Minority Report-style police attention for likely offenders*, Chicago Crime Library (Feb. 24, 2014), **http://www.crimelibrary.com/blog/2014/02/24/heat-list-brings-minority-report-style-police-attention-for-likely-offenders-in-chicago/index.html**.

[61] Mark Guarino, *Can Math Stop Murder?*, Christian Science Monitor (Jul. 20, 2014), **http://www.csmonitor.com/USA/2014/0720/Can-math-stop-murder-video**.

[62] Matt Stroud, *The Minority Report: Chicago's New Police Computer Predicts Crimes, But is it Racist?*, The Verge (Feb. 19, 2014), **http://www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist**.

[63] *Prison Breakthrough*, The Economist (Apr 19, 2014), **http://www.economist.com/news/united-states/21601009-big-data-can-help-states-decide-whom-release-prison-prison-breakthrough** ("Four-fifths of parole boards now use 'risk-assessment' software) technology, says Joan Petersilia of Stanford University.").

[64] *Don't Even Think About It*, The Economist (Jul. 18, 2013), **http://www.economist.com/news/briefing/21582042-it-getting-easier-foresee-wrongdoing-and-spot-likely-wrongdoers-dont-even-think-about-it**.

[65] William Harms, *City Cites Crime Lab Data in Funding Innovative Youth Program*, UChicago News (Feb. 7, 2013), **http://news.uchicago.edu/article/2013/02/07/city-cites-crime-lab-data-funding-innovative-youth-program**.

[66] Aaron Rieke, *Seattle Powers Down Police Wi-Fi Network, Making Room for Public Debate*, Equal Future (Nov. 20, 2013), **http://equalfuture.us/2013/11/20/seattle-wifi**.

[67] Kim Zetter, *Police Contract With Spy Tool Maker Prohibits Talking About Device's Use*, Wired (Mar. 4, 2014), **http://www.wired.com/2014/03/harris-stingray-nda**.

[68] Letter from P.O. Cronin, supra note 59.

[69] Nathan Freed Wessler, *U.S. Marshals Seize Local Cops' Cell Phone Tracking Files in Extraordinary Attempt to Keep Information*, Free Future (Jun. 3, 2014), **https://www.aclu.org/blog/national-security-technology-and-liberty/us-marshals-seize-local-cops-cell-phone-tracking-files**.

[70] American Civil Liberties Union, *You Are Being Tracked: How License Plate Readers Are Being Used to Record Americans' Movements* (Jul. 17, 2013), **https://www.aclu.org/alpr**.

[71] *Id*.

[72] *Id*.

[73] Zetter, *supra* note 67.

[74] *Id*.

[75] Aaron Rieke, *Police Accept Gag Rules on New Surveillance Tech*, Equal Future (Mar. 12, 2014), **http://equalfuture.us/2014/03/12/police-accept-gag-rules-on-new-surveillance-tech**.

[76] Wessler, *supra* note 69.

[77] Maria Kayanan, *Internal Police Emails Show Efforts to Hide Use of Cell Phone Tracking*, Free Future (Jun. 19, 2014), **https://www.aclu.org/blog/national-security-technology-and-liberty/internal-police-emails-show-efforts-hide-use-cell**.

[78] See U.S. Department of Homeland Security, *CCTV: Developing Privacy Best Practices* (Dec. 2007), **http://www.dhs.gov/xlibrary/assets/privacy/privacy_rpt_cctv_2007.pdf**.

[79] David Robinson, *Police Departments Quickly Adopting "Body-Worn Cameras" That Put Citizen Encounters On Video*, Equal Future (Sep. 25, 2013), **http://equalfuture.us/2013/09/25/police-departments-quickly-adopting-body-worn-cameras-that-put-citizen-encounters-on-video**.

[80] Christopher Mims, *What Happens When Police Officers Wear Body Cameras*, Wall St. J. (Aug. 18, 2014), **http://online.wsj.com/articles/what-happens-when-police-officers-wear-body-cameras-1408320244**

[81] Letter from Sherrilyn Ifill, President, NAACP Legal Defense & Education Fund, to Eric Holder (Aug. 14, 2014), **http://www.naacpldf.org/files/case_issue/8-14-2014%20Letter%20to%20AG%20Holder%20re%20use%20of%20excessive%20force%20by%20police.pdf** ("Consistent with its financial and practical influence over state and local law enforcement agencies, the DOJ should promote the use of body-worn video cameras. Properly obtained video evidence produces an objective account of interactions between police and citizens. This improves the accuracy of investigations into police brutality and misconduct, contextualizes citizen encounters with the police, provides training opportunities to officers about appropriate police practices, and serves as an independent check on police conduct.").

[82] Azi Paybarah, *City maintains not-quite-yet position on NYPD*, Capital (Aug. 6, 2014), cameras **http://www.capitalnewyork.com/article/city-hall/2014/08/8550154/city-maintains-not-quite-yet-position-nypd-cameras**.

[83] J. David Goodman, *New York Police Officers to Start Using Body Cameras in a Pilot Program*, N.Y. Times (Sep. 4, 2014), **http://www.nytimes.com/2014/09/05/nyregion/new-york-police-officers-to-begin-wearing-body-cameras-in-pilot-program.html**.

[84] Robinson *supra* note 79.

[85] Paybarah *supra* note 82.

[86] Marc Santora, *Order That Police Wear Cameras Stirs Unexpected Reactions*, N.Y. Times (Aug. 13, 2013), **http://www.nytimes.com/2013/08/14/nyregion/order-that-police-wear-cameras-stirs-unexpected-reactions.html**.

[87] *Id*.

[88] Sarah Lai Stirland, *High-Tech NYPD Is Body Camera Shy, As Departments Across Country Embrace Them*, TechPresident (Sep. 23, 2013), **http://techpresident.com/news/24357/bloomberg-administration-fights-body-cameras-police-departments-across-country-embrace**.

[89] Police Executive Research Forum, *Police Leaders Explore Growing Use of Body Cameras At PERF Town Hall Meeting in Philadelphia* (Sep. 2013), **http://www.policeforum.org/assets/docs/Subject_to_Debate/Debate2013/debate_2013_sepoct.pdf**.

[90] Santora, *supra* note 86.

[91] Stirland, *supra* note 88.

[92] *Id*.

[93] EFE, *U.S. Mulls Over Putting Cameras On Border Patrol Agents*, Fox News Latino (Mar. 26, 2014), **http://latino.foxnews.com/latino/news/2014/03/25/us-considers-placing-cameras-on-border-patrol-agents**.

[94]Jay Stanley, *Police Body-Mounted Cameras: With Right Policies in Place, a Win For All*, ACLU (Oct. 2013), **https://www.aclu.org/files/assets/police_body-mounted_cameras.pdf**.

[95] Sara Libby, *Even When Police Do Wear Cameras, Don't Count on Seeing the Footage*, City Lab (Aug 18., 2014), **http://www.citylab.com/crime/2014/08/even-when-police-do-wear-cameras-you-cant-count-on-ever-seeing-the-footage/378690**.

[96] Scott Shane and Colin Moynihan, *Drug Agents Use Vast Phone Trove, Eclipsing N.S.A.'s*, N.Y. Times (Sep. 1, 2013), **http://www.nytimes.com/2013/09/02/us/drug-agents-use-vast-phone-trove-eclipsing-nsas.html**.

[97] *Id*.

[98] Office of National Drug Control Policy, *Los Angeles Hemisphere: Law Enforcement Sensitive*, at 5, **http://www.nytimes.com/interactive/2013/09/02/us/hemisphere-project.html**.

[99] *Id*. at 12.

[100] John Shiffman and Krista Cooke, *Exclusive: U.S. Directs Agents To Cover Up Program Used To Investigate Americans*, Reuters (Aug. 5, 2013), **http://www.reuters.com/article/2013/08/05/us-dea-sod-idUSBRE97409R20130805**.

[101] *Id*.

[102] Karen McVeigh, *US drug agency surveillance unit to be investigated by Department of Justice*, Guardian (Aug. 6, 2013), **http://www.theguardian.com/world/2013/aug/06/justice-department-surveillance-dea**.

[103] The Washington Post, *The Top Secret Network of the Government and it's Contractors* (Sep. 2010), **http://projects.washingtonpost.com/top-secret-america/network/#/overall/most-activity**.

[104] The Washington Post, *NSA slides explain the PRISM data-collection program* (Jul. 10, 2013), **http://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents**.

[105] Jessica Lussenhop, *Is Anne Marie Rasmusson too hot to have a driver's license?*, City Pages (Feb. 22, 2012), **http://www.citypages.com/2012-02-22/news/is-anne-marie-rasmusson-too-hot-to-have-a-driver-s-license**.

[106] Lauren Sher, *Cop Issues Speeding Ticket, Asks Driver for a Date and She Sues Him*, ABC News Blogs (Jan. 5, 2012), **http://news.yahoo.com/blogs/abc-blogs/cop-issues-speeding-ticket-asks-driver-date-she-002427538.html**.

[107] John Barna, *Voorhees Township cop suspended on allegations he looked up woman's license information, 'friended' her on Facebook*, Glouster County Times (Jul. 24, 2012), **http://www.nj.com/gloucester-county/index.ssf/2012/07/voorhees_township_officer_susp.html**.

[108] Dareh Gregorian & Ginger Adams Otis, *NYPD detective hacked into computers to get fellow officers' email, cell phone info: Feds*, N.Y. Daily News (May. 21, 2013), **http://www.nydailynews.com/news/crime/feds-nypd-detective-hacked-computers-article-1.1350359**.

[109] *NC policeman misused database to stalk woman, authorities say*, WNCT News (Nov. 26, 2013), **http://www.wbtw.com/story/24078329/nc-policeman-charged-with-stalking-a-woman**.

[110] *See* Sher, *supra* note 106.

[111] *See* Gregorian & Otis, *supra* note 108.

[112] Frank Washkuch Jr., *Conn. police sergeant charged with computer crime*, S.C. Magazine (Feb. 7, 2008), **http://www.scmagazine.com/conn-police-sergeant-charged-with-computer-crime/article/105085/**

[113] Jerzy Shedlock, *'Boneheaded' former Anchorage officer sentenced for misuse of police database*, Alaska Dispatch News (Apr. 11, 2014), **http://www.alaskadispatch.com/article/20140411/boneheaded-former-anchorage-officer-sentenced-misuse-police-database**.

[114] Cyrus Farivar, *Snowden: NSA employees routinely pass around intercepted nude photos*, Ars Technica (Jul. 17, 2014), **http://arstechnica.com/tech-policy/2014/07/snowden-nsa-employees-routinely-pass-around-intercepted-nude-photos**.

[115] Brian Ross, Vic Walter & Anna Schecter, *Exclusive: Inside Account of U.S. Eavesdropping on Americans*, ABC News (Oct. 9, 2008), **http://abcnews.go.com/Blotter/exclusive-inside-account-us-eavesdropping-americans/story?id=5987804**.

[116] Letter from Dr. George Ellard, Inspector General, National Security Agency, to Senator Charles E. Grassley (Sep. 11, 2013), **http://s3.documentcloud.org/documents/799762/nsa-surveillance-09-11-13-response-from-ig-to.pdf**.

[117] Timothy B. Lee, *5 Americans who used NSA facilities to spy on lovers*, The Switch—The Washington Post (Sep. 27, 2013), **http://www.washingtonpost.com/blogs/the-switch/wp/2013/09/27/5-americans-who-used-nsa-facilities-to-spy-on-lovers**.

[118] Charlie Savage, *N.S.A. Calls Violations of Privacy 'Minuscule'*, N.Y. Times (Aug. 16, 2013), **http://www.nytimes.com/2013/08/17/us/nsa-calls-violations-of-privacy-minuscule.html** ("The report showed about 100 errors by analysts in making queries of databases of already-collected communications data; by comparison, he said, the agency performs about 20 million such queries each month.").

[119] Siobhan Gorman, *NSA Officers Spy on Love Interests*, Wall St. J. (Aug. 23, 2013), **http://blogs.wsj.com/washwire/2013/08/23/nsa-officers-sometimes-spy-on-love-interests**.

[120] U.S. Census Bureau, *United States Census 2010*, **https://www.census.gov/schools/pdf/2010form_info.pdf**.

[121] U.S. Census Bureau, *1940 Overview*, **https://www.census.gov/history/www/through_the_decades/overview/1940.html**.

[122] U.S. Census Bureau, *American Community Survey – History*, **https://www.census.gov/history/www/programs/demographic/american_community_survey.html**.

[123] NAACP, *Census Fact Sheet*, **http://www.naacp.org/pages/census-fact-sheet**.

[124] State of New Jersey, *Department of Labor and Workforce Development, 50 Ways Census Data Are Used*, **http://lwd.dol.state.nj.us/labor/lpa/census/2010/50WaysDataUsed.html**.

[125] United States Government Accountability Office, *2010 Census: Census Bureau Should Take Action to Improve the Credibility and Accuracy of Its Cost Estimate for the Decennial Census* (Jun. 2008), **http://www.gao.gov/assets/280/276782.pdf** (The cost of the census in inflation-adjusted dollars has roughly doubled each decade since 1970 from $1 billion in 1970 to $14 billion in 2010, a figure not completely accounted for by the rising number of housing units.).

[126] John H. Thompson, *Ensuring an Accurate and Affordable 2020 Census, Oral Testimony Before the Subcommittee on Federal Workforce, U.S. Postal Service and the Census* (Sep. 11, 2013), **https://www.census.gov/newsroom/releases/pdf/09112013_thompson_statement.pdf**.

[127] National Research Council, *A Census that Mirrors America: Interim Report* 30 (1993), **http://www.nap.edu/openbook.php?record_id=2234**.