

## Executive Summary

Data & Civil Rights  
October 30, 2014 – Washington, D.C.  
<http://www.datacivilrights.org/>

*This document represents a general summary of the discussions that took place at the Data & Civil Rights gathering. Not all attendees were involved in every part of the conversation, nor does this document necessarily reflect the views and beliefs of individual attendees.*

### Overview

On October 30, 2014, the Data & Society Research Institute, The Leadership Conference on Civil and Human Rights, and New America’s Open Technology Institute co-hosted an event entitled “Data & Civil Rights.” The aim of the event was to convene civil rights leaders, technologists, and key representatives from industry and government in and across the sectors of Education, Consumer Finance, Health, Criminal Justice, Housing, and Employment to identify and examine how the “big data” phenomenon affects civil rights issues. The conversations ranged from the consideration of how technology perpetuates historical discrimination to the possibilities of using data collection to serve civil rights goals, particularly by measuring inequalities or disparities.

The convening paid considerable attention to the historical dynamics of civil rights and the technical realities of the “big data” phenomenon. The event had three main narratives: (1) the roots and contemporary state of civil rights issues, which centered primarily on discrimination on the basis of protected classes, and issues of privacy; (2) the inner workings of the technology and how and when it can create discriminatory outcomes and impacts, particularly through algorithmic decision-making; and (3) the next steps for these discussions, especially in the areas of policy-making, government actions, technology development, generating social change, industry innovation, and new research.

Given the wide range of constituents in the room, the event sought to develop a common language among a group of people broadly interested in issues of civil rights, equity, and data-driven technology. Participants included those who were unfamiliar with the ways in which technology is transforming the civil rights landscape, as well as those unfamiliar with the social and civil rights implications of the technologies that are being designed and used. These gaps in knowledge raised an opportunity to address key questions: How can civil rights values be embedded into the designs of emerging technologies? Is it possible to develop a new civil rights science that aids the cause of equal opportunity and social justice?

Throughout the day, participants were challenged to explain and contextualize civil rights issues and to clarify the goals of civil rights advocacy, and conversely, to explain the capabilities, limits, and pitfalls of technology. Discussions also focused on what it takes to achieve, and how to measure, equality and equity. For example, all schools might receive equal funding, but if one school is teaching students who require more resources, like English language learners, than another, than they require greater inputs/resources to achieve an equal outcome with schools

whose students require less resources. In addition, discussion surfaced another area of tension regarding what is fair and what is lawful: many kinds of unscrupulous uses of data may not be unlawful, even their effects are unfair. To the surprise of many, some of the most insidious examples of “big data” applications or outcomes represent unintentional discrimination— byproducts of technical designs and practices that were never intended to cause harm and are hard to prevent from happening again.

These discoveries provoked discussion about the adequacy of existing civil rights protections and the ways that new technologies threaten to undermine or support them. A recurrent theme centered on the need for new or expanded uses of existing laws and regulations, particularly the Fair Credit and Reporting Act, to protect civil rights across sectors in response to, and in anticipation of, the flood of data. Many participants shared a common desire to think about using data-driven technologies to improve socio-economic mobility for historically marginalized groups and reduce widening gaps in equity. In this sense, the convening emphasized the possibility for technology to optimize for civil rights goals, above and beyond what the law provides. Many questioned whether legislation and policy work can keep pace with new technologies, and whether technology can offer a more efficient avenue than laws for designing and implementing systems that perpetuate civil rights values.

Across all of the workshop discussions, and as participants grappled with the ways to address the ways that different actors to use data and the impact of data-driven, algorithmic systems, several themes emerged: the right to dispute incorrect or inaccurate data; whether further regulatory measures were needed; the issue of auditing algorithmic systems, and how to address bad results that come from neutral or malicious intentions; and the best ways to measure the impact of data-driven technologies.

As the day progressed, many conversations about data and fairness pivoted around the notions of transparency and accountability, in relation to both consumers and regulators. How can we achieve transparency about data-driven systems in ways that are meaningful to the people who are affected by them, without expressly requiring the companies that use data-driven technologies to disclose proprietary information? Often, these discussions centered on the legal definition of discrimination and legal remedies. For examples, if “big data” systems produce decisions that have a disparate impact effect on protected classes in areas where disparate impact theory applies, system creators, owners, and operators challenged by a disparate impact lawsuit have a legal obligation to explain how their systems function or how they’ve taken steps to reduce discriminatory effects, in order to prevail. How do you require or incentivize disclosures about algorithmic decisions, or incentivize companies to perform unit tests on the key attributes (protected class characteristics) that we care about? If there is another way to build the mousetrap that still gets the job done, how do we incentivize that process, especially at the design stage, rather than as a retrofit, or as the outcome of legal challenges?

As workshop sessions revealed common themes and challenges, the need for more refined analyses of the problems and prospective sector-specific solutions became evident. “Big data” does not function as the source of, or solution to, harms in the same way in each sector. Participants raised the importance of identifying differences between different policy areas, particularly regulatory regimes that involve public monopolies (like the Criminal Justice system), public monopolies threatened by market forces (like EdTech developments in Education), mixed public and private monopolies (like Health Care), or a marketplace of private companies (like

Consumer Finance). In addition, each sector is at a different stage of development or reliance on data-driven technologies and represents varying levels of receptivity to data regulation.

By and large, participants focused on asking difficult questions rather than proposing concrete solutions. This interrogatory tendency stemmed from the absence of adequate empirical data to ground claims and concerns. Participants also raised and debated the lack of concrete harms: some highlighted the difficulty of identifying harms given the nature of discrimination and others suggested that part of what's at stake here stems from the fundamental difference in types of negative consequences that may not look like harms in a traditional legal sense. In conversations about positive uses of technology, speculation trumped evidence, prompting critics to be wary of the claims made by those seeking to use technology for good. Above all, these discussions laid bare a critical need for more concrete information about the positive and negative outcomes of using "big data" technologies in civil rights contexts.

Participants were encouraged to use the experience they had at the conference to educate not just the people in the room, but to take the conversations they had to the public. One of the key take-aways from the conference overall is the urgency and importance of broadening and deepening conversations and developing leadership from the mixed group of stakeholders who can represent and engage the public, as well as corporate and industrial actors.

As we continue to reflect on the discussions that took place during and around this event, we will continue to work across sectors to envision a path forward. As you read the materials from the event, please feel free to contact us. We are open to feedback and would love to hear suggestions about additional next steps.

*danah boyd, Seeta Peña Gangadharan, and Corrine Yu*