

Breakout Discussion Notes: Technology Development

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

This document was produced based on notes taken during the Technology Development breakout session of the Data & Civil Rights conference. This document represents a general summary of the discussion that took place. Not all attendees were involved in every part of the conversation, nor does this document necessarily reflect the views and beliefs of individual attendees. For an overview of the breakout sessions, including a description of the questions participants were asked to consider, see: <http://www.datacivilrights.org/pubs/2014-1030/BreakoutOverview.pdf>

Overview

The group in this session focused on the development of new technologies in the context of civil rights. How can new “big data” technologies be used to identify civil rights violations, and can those same technologies be used to intervene in and improve areas of social inequality? The group tried to balance and reconcile the differing goals and values of various invested actors—from the developers and owners of new data-producing services, to government regulators, to individual members of vulnerable populations. Many “big data” techniques seem to produce a natural tension between efficacy and fairness. Such an observation prompts a re-consideration of some central definitions: What is fairness? What is justice? What is transparency? By wrestling with and proposing some answers to these definitional questions, the group was able to point to some possible sets of future best practices for the handlers of data and imagine some starting points on the road to better legal regulation.

Themes and Discussion Topics

The technological developments labeled under the umbrella of “big data” seem to have a deeply ambivalent relationship to the reality of civil rights. It is an exciting possibility that new datasets and data techniques might be capable of identifying previously invisible civil rights violations as well as enable new interventions for raising the level of civil equality. At the same time, however, the very same datasets and techniques can just as easily be used to further the violations of civil liberties, or create new spaces of inequality. In any discussion centered on civil rights and big data, there must be an acknowledgement of the tension between the *benefits* and *harm* capable of being produced by a particular technological arrangement or design.

Participants discussed an example of this technological reality taken from the sphere of healthcare: in particular, they highlighted a case where Asian patients are systematically categorized according to 39 different categories of data, which is specifically useful because there are genetic variants, for example, that stratify with more granularity than ‘pan-Asian’, which affect treatment protocols. This level of specificity is extremely effective *medically*—it allows for early diagnosis and more effective treatments—but the same details, repurposed, could also be used to render significant economic and social harm—serving as the basis for the refusal of

employment, insurance, or otherwise enable new forms of discrimination.

The participants acknowledged that many of the questions raised by technological developments were the same as those raised in other sessions. That is, technological development cannot be taken as a wholly separate sphere of concern, but must be integrated with other areas and topics. This leads to many questions. What sorts of data models might prove useful for different areas of knowledge? What datasets would one need to regulate health care? What datasets would help monitor civil rights violations? What datasets would be useful for criminal justice? And in each case, what are the different techniques (categorical, analytical, predictive) that are most effective? Finally, when choosing a dataset and/or suite of data techniques for a particular application, how can we identify when either (or both) produces bias? How can we assess, compare, and/or avoid bias in data?

In part, the difficulties in answering the above questions come from the range of applications to which any new technology can be applied. Even technologies designed with the best of intentions can open up the space for new abuses or inadvertent inequalities. One must consider issues of first and second order interpretations when evaluating the impact of a new technology. Any time one set of users has more power or access than others, the system risks enabling unfair behavior. In addition to accurately identifying the effects of technologies on conditions of social fairness, it also becomes important to interrogate any assumed definitions of fairness. What is fairness and how can it be assessed and tested for? Is there a system of accountability or auditing that could be used to test for biases against minority users or inputs of these datasets?

For technologists to implement values into code, it becomes imperative to cleanly define the values in an implementable way. In short, technologists need to know what they are solving for and that is not always clear in the conversation about data and civil rights. Even basic terms are not consistently or coherently defined. Technologists need clear boundaries, clear test cases, and mechanisms of assessment. All of this is missing in the murky conversation about misuse of data.

Because power can become unequally distributed in “big data” systems, it is important to inquire into who is designing these systems, and where they’re being installed. Who will be the individuals or groups that will use these systems to monitor the conditions relevant to civil rights? The importance of these questions implies that the design of such systems should strive for *transparency* in order to ensure equity. But (once again) what do we mean by transparency? Transparency does not manifest in just a single way. There are four dimensions of transparency by which new technologies should be evaluated:

- *Data Availability*: Does the public know what data are being used/analyzed?
- *Provenance*: Does the public know where the data come from?
- *User Understanding*: Does the public understand the processes applied to the data?
- *User Awareness*: What portion of the effected public is aware of the underlying process? What constitutes broad awareness?

The need for transparency naturally dovetails with the importance of accountability. If transparent operations reveal some condition of inequality, there must then be a mechanism in place to host a public discussion of the issue and translate consensus into new effective regulation. This means creating processes where regulating agencies, technologists, lawyers, and individual citizens can work together to map decisions onto legal structures. Many of the issues raised by new technologies require solutions different from their historical precursors, and it takes a specific

technical knowledge to recognize and respond to this. This has contributed to a situation where many policy decisions are flowing from technologists (Silicon Valley) to regulators (Washington DC). On the one hand, this gives technologists an opportunity to positively influence policy through the proactive adoption of fair technologies and practices. On the other hand, political decision-making should be an essentially *collaborative process*, and therefore must strive to include implicated parties from both within and without these technology companies.

Some other key issues that emerged include:

- *Agency*: Do users have the agency to assert control over a technology? Do the legal mechanisms of protection actually empower users in navigating global technical information flows?
- *Data Flows*: The designers technical systems and original collectors of data may have great intentions, but what happens when technologies are used in unintended ways or when third party access to the data is where trouble begins?
- *Privacy as an Elite Frame*: Many of those who speak loudest about privacy in technology aren't those who are most marginalized. This is the frame for thinking about data control that is most common in technical circles, but what other frames would be more likely to mean something to disenfranchised people?

Areas for Further Exploration

Despite some discussion during the panel, the group acknowledged that a more detailed definition of fairness still needed to be developed. Without a concrete description of what constitutes fairness, consistent regulation cannot be created.

In addition, once criteria of fairness are defined, there need to be appropriate techniques for *testing* for that fairness. This will be a process that needs careful design, as gathering the data necessary for such tests could (of course) produce new opportunities for inequality. Therefore test cases must be thoroughly assessed for their representation of all involved parties for users to companies to vendors and beyond.

Technical researchers started questioning whether it was possible for those working with data to make assertions about their practices and for those assertions to be validated technically without revealing the data itself. This, alongside the broader question of technical auditing, is a fruitful area for further technical research.

The question remains, as well, whether making these decisions and arranging these tests based criteria of fairness will actually improve the state of civil rights in the context of new technologies.